

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2002, 10:34:58 ; Search time 81.22 Seconds
(without alignments)
3053.354 Million cell updates/sec

Title: US-09-680-514-8
Perfect score: 1095
Sequence: 1 ATGGAGCTGACTGAATTGCT.....TACGCCACTTCCCGCCGCC 1095

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2.6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2.6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2.6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2.6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2.6/ptodata/2/ina/PTUS_COMB.seq.*
6: /cgn2.6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	518.8	47.4	1059	US-08-471-045-57	Sequence 57, Appl
2	518.8	47.4	1059	US-08-469-712A-57	Sequence 57, Appl
3	518.8	47.4	1062	US-08-347-029-3	Sequence 3, Appl
4	518.8	47.4	1062	US-08-457-254-3	Sequence 3, Appl
5	518.8	47.4	1062	US-08-484-257-18	Sequence 18, Appl
6	518.8	47.4	1062	US-08-482-212-1	Sequence 1, Appl
7	518.8	47.4	1062	US-08-999-927-1	Sequence 1, Appl
8	518.8	47.4	1062	PCT-US94-08806-18	Sequence 18, Appl
9	518.8	47.4	1062	PCT-US95-01829-1	Sequence 1, Appl
10	518.8	47.4	1062	PCT-US95-14932-3	Sequence 3, Appl
11	518.8	47.4	1062	PCT-US95-16626-3	Sequence 3, Appl
12	518.8	47.4	1342	US-08-388-779A-1	Sequence 1, Appl
13	518.8	47.4	1342	US-08-591-070A-1	Sequence 1, Appl
14	518.8	47.4	1342	US-08-413-803-24	Sequence 24, Appl
15	518.8	47.4	1342	US-08-321-488A-24	Sequence 24, Appl
16	518.8	47.4	1342	US-08-414-161B-1	Sequence 1, Appl
17	518.8	47.4	1342	US-08-927-855-1	Sequence 1, Appl
18	518.8	47.4	1342	PCT-US95-03776-24	Sequence 24, Appl
19	518.8	47.4	1342	PCT-US95-03776-26	Sequence 26, Appl
20	518.8	47.4	1795	US-08-330-517-1	Sequence 1, Appl
21	517.2	47.2	605	US-08-388-779A-3	Sequence 3, Appl
22	517.2	47.2	605	US-08-591-070A-3	Sequence 3, Appl
23	517.2	47.2	605	US-08-927-855-3	Sequence 3, Appl
24	486.8	44.5	525	US-08-434-411-1	Sequence 1, Appl
25	486.8	44.5	525	US-08-434-402-1	Sequence 1, Appl
26	486.8	44.5	525	US-08-783-288-1	Sequence 1, Appl
27	486.8	44.5	525	US-08-890-640-1	Sequence 1, Appl

28	486.8	44.5	525	6	5194592-25	Patent No. 5194592
29	486	44.4	921	3	US-08-469-318-60	Sequence 60, Appl
30	486	44.4	921	3	US-08-469-318-87	Sequence 87, Appl
31	486	44.4	921	3	US-08-468-609A-60	Sequence 60, Appl
32	486	44.4	921	3	US-08-468-609A-87	Sequence 87, Appl
33	486	44.4	921	5	PCT-US95-01185-60	Sequence 60, Appl
34	486	44.4	921	5	PCT-US95-01185-87	Sequence 87, Appl
35	482.8	44.1	921	3	US-08-469-318-84	Sequence 84, Appl
36	482.8	44.1	921	3	US-08-468-609A-84	Sequence 84, Appl
37	481.2	43.9	1017	3	PCT-US95-01185-84	Sequence 84, Appl
38	481.2	43.9	1017	3	US-08-469-318-67	Sequence 67, Appl
39	481.2	43.9	1017	3	US-08-468-609A-67	Sequence 67, Appl
40	481.2	43.9	1017	5	PCT-US95-01185-67	Sequence 67, Appl
41	478	43.7	921	3	US-08-469-318-56	Sequence 56, Appl
42	478	43.7	921	3	US-08-469-318-85	Sequence 85, Appl
43	478	43.7	921	3	US-08-468-609A-56	Sequence 56, Appl
44	478	43.7	921	3	US-08-468-609A-85	Sequence 85, Appl
45	478	43.7	921	5	PCT-US95-01185-56	Sequence 56, Appl

ALIGNMENTS

RESULT 1
US-08-471-045-57
; Sequence 57, Application US/08471045
; Patent No. 6060047
; GENERAL INFORMATION:
; APPLICANT: Abrams, Mark A.
; APPLICANT: Bauer, S. C.
; APPLICANT: Braford-Goldberg, Sarah R.
; APPLICANT: Caparon, Mair H.
; APPLICANT: Easton, Alan M.
; APPLICANT: Klein, Barbara K.
; APPLICANT: McKearn, John P.
; APPLICANT: Ollins, Peter O.
; APPLICANT: Paik, Kumnan
; APPLICANT: Thomas, John W.
; TITLE OF INVENTION: Co-administration of Interleukin-3
; TITLE OF INVENTION: Mutant Polypeptides with CSF's for Multi-lineage
; TITLE OF INVENTION: Hematopoietic Cell Production
; NUMBER OF SEQUENCE: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,
; ADDRESS: Corporate Patent Dept.
; STREET: P. O. Box 5110
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60680
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,045
; FILING DATE: 06-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/01184
; FILING DATE: 02-FEB-1995
; APPLICATION NUMBER: US 08/193,373
; FILING DATE: 04-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Dennis A.
; REGISTRATION NUMBER: 34,547
; REFERENCE/DOCKET NUMBER: C-2789/3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (708)470-6501
; TELEFAX: (708)470-6881
; INFORMATION FOR SEQ ID NO: 57:

SEQUENCE CHARACTERISTICS:
LENGTH: 1059 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-471-045-57

Query Match 47.4%; Score 518.8; DB 3; Length 1059;
Best Local Similarity 99.6%; Pred. No. 1.4e-116;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATTCCTCTGCTGGTTCATGCTTCTCTAACTCAAGGCTTAACGCTG 60
DB 1 ATGGAGCTGACTGAATTCCTCTGCTGGTTCATGCTTCTCTAACTCAAGGCTTAACGCTG 60
QY 61 TCCAGCCCGGCTCTCTGCTTGTGACCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
DB 61 TCCAGCCCGGCTCTCTGCTTGTGACCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
QY 121 CATGCTCTTCCACAGCAGTGGCAGTGGCCAGAGGTTTCAACCTTTGCTACACCTGTC 180
DB 121 CATGCTCTTCCACAGCAGTGGCAGTGGCCAGAGGTTTCAACCTTTGCTACACCTGTC 180
QY 181 CTGCTGCTGCTGTGGAGCAGTGGCAGTGGCAGTGGCAGTGGCAGTGGCAGTGGCAG 240
DB 181 CTGCTGCTGCTGTGGAGCAGTGGCAGTGGCAGTGGCAGTGGCAGTGGCAGTGGCAG 240
QY 241 GCACAGACATTTCTGGAGCAGTGGCAGTGGCAGTGGCAGTGGCAGTGGCAGTGGCAG 300
DB 241 GCACAGACATTTCTGGAGCAGTGGCAGTGGCAGTGGCAGTGGCAGTGGCAGTGGCAG 300
QY 301 CAACCTGGGACCCACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
DB 301 CAACCTGGGACCCACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 361 CTCCTTGGGGCCCTCGAGAGCTGGCAGTGGCAGTGGCAGTGGCAGTGGCAGTGGCAG 420
DB 361 CTCCTTGGGGCCCTCGAGAGCTGGCAGTGGCAGTGGCAGTGGCAGTGGCAGTGGCAG 420
QY 421 GCTCACAAGATCCCAATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 421 GCTCACAAGATCCCAATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 481 CGTTTCTGATGCTTGTAGGAGGCTCCACCTCTGCGTAGGG 522
DB 481 CGTTTCTGATGCTTGTAGGAGGCTCCACCTCTGCGTAGGG 522

RESULT 2

US-08-469-712A-57
Sequence 57, Application US/08469712A
Patent No. 6093395

GENERAL INFORMATION:

APPLICANT: Abrams, Mark A.
APPLICANT: Bauer, S. C.
APPLICANT: Braford-Goldberg, Sarah R.
APPLICANT: Caparon, Maïre H.
APPLICANT: Easton, Alan M.
APPLICANT: Klein, Barbara K.
APPLICANT: McKearn, John P.
APPLICANT: Orlins, Peter O.
APPLICANT: Paik, Kumnan
APPLICANT: Thomas, John W.
TITLE OF INVENTION: Co-administration of Interleukin-3
TITLE OF INVENTION: Mutant Polypeptides with CSF's for Multi-lineage
TITLE OF INVENTION: Hematopoietic Cell Production
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,
ADDRESSEE: Corporate Patent Dept.
STREET: P. O. Box 5110

CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60680
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,712A
FILING DATE: 06-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/01184
FILING DATE: 02-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/193,373
FILING DATE: 04-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bennett, Dennis A.
REGISTRATION NUMBER: 34,547
REFERENCE/DOCKET NUMBER: C-2789/2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (708)470-6501
TELEFAX: (708)470-6881
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 1059 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-469-712A-57

Query Match 47.4%; Score 518.8; DB 3; Length 1059;

Best Local Similarity 99.6%; Pred. No. 1.4e-116;

Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATTCCTCTGCTGGTTCATGCTTCTCTAACTCAAGGCTTAACGCTG 60
DB 1 ATGGAGCTGACTGAATTCCTCTGCTGGTTCATGCTTCTCTAACTCAAGGCTTAACGCTG 60
QY 61 TCCAGCCCGGCTCTCTGCTTGTGACCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
DB 61 TCCAGCCCGGCTCTCTGCTTGTGACCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
QY 121 CATGCTCTTCCACAGCAGACTGAGCCAGTGGCCAGAGGTTTCAACCTTTGCTACACCTGTC 180
DB 121 CATGCTCTTCCACAGCAGACTGAGCCAGTGGCCAGAGGTTTCAACCTTTGCTACACCTGTC 180
QY 181 CTGCTGCTGCTGTGGACTTTAGCTTGGAGAGTGGAAACCCAGATGGAGGAGACCAAG 240
DB 181 CTGCTGCTGCTGTGGACTTTAGCTTGGAGAGTGGAAACCCAGATGGAGGAGACCAAG 240
QY 241 GCACAGGACATTTCTGGGAGCAGTGGCCTTCTGCTGGAGGAGTGGCAGCAGCGGGA 300
DB 241 GCACAGGACATTTCTGGGAGCAGTGGCCTTCTGCTGGAGGAGTGGCAGCAGCGGGA 300
QY 301 CAACCTGGGACCCACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
DB 301 CAACCTGGGACCCACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 361 CTCCTTGGGGCCCTCGAGAGCTGGCAGTGGCAGTGGCAGTGGCAGTGGCAGTGGCAG 420
DB 361 CTCCTTGGGGCCCTCGAGAGCTGGCAGTGGCAGTGGCAGTGGCAGTGGCAGTGGCAG 420
QY 421 GCTCACAAGATCCCAATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 421 GCTCACAAGATCCCAATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 481 CGTTTCTGATGCTTGTAGGAGGCTCCACCTCTGCGTAGGG 522

Qy 1 ATGAGCTGACTGAATTGCTCCTCGTGGTCATGCTTCTCCTAACTGCAAGGCTAACGCTG 60
|||||
Db 1 ATGAGCTGACTGAATTGCTCCTCGTGGTCATGCTTCTCCTAACTGCAAGGCTAACGCTG 60

Qy	61	TCAGCCCGGCTCCTCGCTGTGTGACCTCCGAGTCCCTCAGTAAACTGCTTCGTGACTCG	120
Db	61	TCAGCCCGGCTCCTCGCTGTGTGACCTCCGAGTCCCTCAGTAACTGCTTCGTGACTCC	120
Qy	121	CATGTCCCTTCACGACGACTGAGCCAGTGCACAGAGTTACCCCTTTGCTCTACACCTGTG	180
Db	121	CATGTCCCTTCACGACGACTGAGCCAGTGCACAGAGTTACCCCTTTGCTCTACACCTGTG	180
Qy	181	CTGCTGCCTGTGTGGACTTTAGCTTTGGGAGAATGGAAACCCAGATGGAGGAGACCAAG	240
Db	181	CTGCTGCCTGTGTGGACTTTAGCTTTGGGAGAATGGAAACCCAGATGGAGGAGACCAAG	240
Qy	241	GCACAGGACATCTTGGAGCAGTGAACCTTCTGTGGAGGAGTGATGGCAGCACGGGA	300
Db	241	GCACAGGACATCTTGGAGCAGTGAACCTTCTGTGGAGGAGTGATGGCAGCACGGGA	300
Qy	301	CAACTGGGACCCACTTGCCTCTCATCCCTCTGGGGCAGCTTTCTGCACAGTGCCTCTC	360
Db	301	CAACTGGGACCCACTTGCCTCTCATCCCTCTGGGGCAGCTTTCTGCACAGTGCCTCTC	360
Qy	361	CTCCTTTGGGGCCCTGCAGAGCTCCTTGGAACCCAGCTTCTCCACAGGGCAGGACCACA	420
Db	361	CTCCTTTGGGGCCCTGCAGAGCTCCTTGGAACCCAGCTTCTCCACAGGGCAGGACCACA	420
Qy	421	GCTCACAGGATCCCAATGCGCATCTCTGTGAGCTTCCAAACACCTGCTCCGAGGAAGGTG	480
Db	421	GCTCACAGGATCCCAATGCGCATCTCTGTGAGCTTCCAAACACCTGCTCCGAGGAAGGTG	480
Qy	481	CGTTTCTGTATGTTGTAGGAGGTCCACCTCTGCGTACGG	522
Db	481	CGTTTCTGTATGTTGTAGGAGGTCCACCTCTGCGTACGG	522

RESULT 5
US-08-484-257-18
: Sequence 18, Application US/08484257
: Patent No. 5989537
: GENERAL INFORMATION:
: APPLICANT: Holly, Richard D.
: APPLICANT: Lok, Si
: APPLICANT: Foster, Donald C.
: APPLICANT: Hagen, Frederick S.
: APPLICANT: Kaushansky, Kenneth
: APPLICANT: Kuipper, Joseph L.
: APPLICANT: Lofton-Day, Catherine E.
: APPLICANT: Oort, Pieter J.
: TITLE OF INVENTION: Methods for Stimulating Granulocyte
: NUMBER OF SEQUENCES: 21
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: ZymoGenetics, Inc.
: STREET: 1201 Eastlake E.
: CITY: Seattle
: STATE: WA
: COUNTRY: USA
: ZIP: 98102
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/484,257
: FILING DATE:
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Parker, Gary E
: REGISTRATION NUMBER: 31-648
: REFERENCE/DOCKET NUMBER: 95-09
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 206-442-6673
: TELEFAX: 206-442-6678
: INFORMATION FOR SEQ ID NO: 18:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1062 base pairs
; Type: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1059
; US-08-484-257-18

Query Match          47.4%; Score 518.8; DB 2; Length 1062;
Best Local Similarity 99.6%; Pred. No. 1.4e-116;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps

QY      1  ATGCAGCTGACTGAATTGCTTCCTCGTGGTCAATGCTTCTCCTAACTGCAAGGCTAAGCGTG 60
        |||
Db       1  ATGCAGCTGACTGAATTGCTTCCTCGTGGTCAATGCTTCTCCTAACTGCAAGGCTAAGCGTG 60
        |||

QY      61  TCCAGCCCGGCTCCTCGCTGCTTGACCTCCGAGTCTCTCAGTAACTGCTTCTGTAAGCTCC 120
        |||
Db       61  TCCAGCCCGGCTCCTCGCTGCTTGACCTCCGAGTCTCTCAGTAACTGCTTCTGTAAGCTCC 120
        |||

QY      121  CATGTCTTCCACAGCAGACTGAGCCAGTGCACAGAGTTTACCTTTGGCTATACACCTGTG 180
        |||
Db       121  CATGTCTTCCACAGCAGACTGAGCCAGTGCACAGAGTTTACCTTTGGCTATACACCTGTG 180
        |||

QY      181  CTGCTGCCTGCTGTGACTTTAGCTTTGGGAGATGAAACCCAGATGAGGAGACCAAG 240
        |||
Db       181  CTGCTGCCTGCTGTGACTTTAGCTTTGGGAGATGAAACCCAGATGAGGAGACCAAG 240
        |||

QY      241  GCACAGSACATTCTGGGAGCAGTGCACCTTCTGCTGGAGGAGTATGCACACCGGGGA 300
        |||
Db       241  GCACAGSACATTCTGGGAGCAGTGCACCTTCTGCTGGAGGAGTATGCACACCGGGGA 300
        |||

QY      301  CAACTGGGACCCACTTTCCTCTCAATCCCTTCTGGGAGCTTTCTTGGACAGGTCCGCTCTC 360
        |||
Db       301  CAACTGGGACCCACTTTCCTCTCAATCCCTTCTGGGAGCTTTCTTGGACAGGTCCGCTCTC 360
        |||

QY      361  CTCCTTTGGGGCCCTGCAGAGCTTCCTTGAAACCCAGCTTCTTCCACAGGGCAGGACCACA 420
        |||
Db       361  CTCCTTTGGGGCCCTGCAGAGCTTCCTTGAAACCCAGCTTCTTCCACAGGGCAGGACCACA 420
        |||

QY      421  GCTCAAGAGATCCCAATGCCATCTTCTTGAGTCTCCAACTGCTCTCCGAGGAAGGTG 480
        |||
Db       421  GCTCAAGAGATCCCAATGCCATCTTCTTGAGTCTCCAACTGCTCTCCGAGGAAGGTG 480
        |||

QY      481  CGTTTCTCTGATGCTGTAGGAGGGTCCACCCCTCTCGGTACGG 522
        |||
Db       481  CGTTTCTCTGATGCTGTAGGAGGGTCCACCCCTCTCGGTACGG 522
        |||

```

```

RESULT      6
US-08-482-212-1
; Sequence 1, Application US/08482212
; Patent No. 6013067
; GENERAL INFORMATION:
; APPLICANT: Fibbe, Willem E.
; APPLICANT: Grossmann, Angelika
; TITLE OF INVENTION: METHODS FOR INCREASING HEMATOPOIETIC CELLS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```


SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,212
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Parker, Gary E
REGISTRATION NUMBER: 31-648
REFERENCE/DOCKET NUMBER: 95-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-442-6600 ext 6673
TELEFAX: 206-442-6678
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1062 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1059
US-08-482-212-1

Query Match 47.4%; Score 518.8; DB 3; Length 1062;
Best Local Similarity 99.6%; Pred. No. 1.4e-116;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATGCTCGTGGTCAATGCTTCTCTAACTCAAGGCTAAGCGTG 60
DB 1 ATGGAGCTGACTGAATGCTCGTGGTCAATGCTTCTCTAACTCAAGGCTAAGCGTG 60
QY 61 TCCAGCCCGGCTCCTCTGCTGACCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
DB 61 TCCAGCCCGGCTCCTCTGCTGACCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
QY 121 CATGCTCTTACACAGACTGACCCAGTCCGAGGTTTCCAGTAACTGCTTCGTGACTCC 180
DB 121 CATGCTCTTACACAGACTGACCCAGTCCGAGGTTTCCAGTAACTGCTTCGTGACTCC 180
QY 181 CTGCTGCCCTGCTGGAGCTTTAGCTTGGGAGATGGAACCCAGATGGAGGAGACCAAG 240
DB 181 CTGCTGCCCTGCTGGAGCTTTAGCTTGGGAGATGGAACCCAGATGGAGGAGACCAAG 240
QY 241 GCACAGGACATCTGGAGCAGTGCCTTCTGCTGGAGGAGTGTGACAGCAGCGGGA 300
DB 241 GCACAGGACATCTGGAGCAGTGCCTTCTGCTGGAGGAGTGTGACAGCAGCGGGA 300
QY 301 CAACCTGGGACCCACTTGGAGCAGTGCCTTCTGCTGGAGGAGTGTGACAGCAGCGGGA 360
DB 301 CAACCTGGGACCCACTTGGAGCAGTGCCTTCTGCTGGAGGAGTGTGACAGCAGCGGGA 360
QY 361 CTGCTGGGGCCCTGACAGACCTTCTTGGAACTCCAGTCCCTCCAGGGCAGGACCA 420
DB 361 CTGCTGGGGCCCTGACAGACCTTCTTGGAACTCCAGTCCCTCCAGGGCAGGACCA 420
QY 421 GCTCAAGGATCCCAATGCACTTCTGAGCTTCCACACTGCTCCGAGGAAAGGTG 480
DB 421 GCTCAAGGATCCCAATGCACTTCTGAGCTTCCACACTGCTCCGAGGAAAGGTG 480
QY 481 CGTTTCTGATGCTTAGGAGGCTCCACCTCTGCTACGG 522
DB 481 CGTTTCTGATGCTTAGGAGGCTCCACCTCTGCTACGG 522

RESULT 7
US-08-999-927-1
Sequence 1, Application US/08999927
Patent No. 6059830
GENERAL INFORMATION:
APPLICANT: Kausanskys, Kenneth
TITLE OF INVENTION: Methods of Stimulating Erythropoiesis

TITLE OF INVENTION: Using Hematopoietic Proteins.
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: ZymoGenetics, Inc.
STREET: 1201 Eastlake Avenue East
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98102
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/999,927
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/347,748
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Parker, Gary E
REGISTRATION NUMBER: 31-648
REFERENCE/DOCKET NUMBER: 94-09C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-442-6673
TELEFAX: 206-442-6678
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1062 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1059
US-08-999-927-1

Query Match 47.4%; Score 518.8; DB 3; Length 1062;
Best Local Similarity 99.6%; Pred. No. 1.4e-116;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATGCTCGTGGTCAATGCTTCTCTAACTCAAGGCTAAGCGTG 60
DB 1 ATGGAGCTGACTGAATGCTCGTGGTCAATGCTTCTCTAACTCAAGGCTAAGCGTG 60
QY 61 TCCAGCCCGGCTCCTCTGCTGACCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
DB 61 TCCAGCCCGGCTCCTCTGCTGACCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
QY 121 CATGCTCTTACACAGACTGACCCAGTCCGAGGTTTCCAGTAACTGCTTCGTGACTCC 180
DB 121 CATGCTCTTACACAGACTGACCCAGTCCGAGGTTTCCAGTAACTGCTTCGTGACTCC 180
QY 181 CTGCTGCCCTGCTGGAGCTTTAGCTTGGGAGATGGAACCCAGATGGAGGAGACCAAG 240
DB 181 CTGCTGCCCTGCTGGAGCTTTAGCTTGGGAGATGGAACCCAGATGGAGGAGACCAAG 240
QY 241 GCACAGGACATCTGGAGCAGTGCCTTCTGCTGGAGGAGTGTGACAGCAGCGGGA 300
DB 241 GCACAGGACATCTGGAGCAGTGCCTTCTGCTGGAGGAGTGTGACAGCAGCGGGA 300
QY 301 CAACCTGGGACCCACTTGGAGCAGTGCCTTCTGCTGGAGGAGTGTGACAGCAGCGGGA 360
DB 301 CAACCTGGGACCCACTTGGAGCAGTGCCTTCTGCTGGAGGAGTGTGACAGCAGCGGGA 360
QY 361 CTGCTGGGGCCCTGACAGACCTTCTTGGAACTCCAGTCCCTCCAGGGCAGGACCA 420
DB 361 CTGCTGGGGCCCTGACAGACCTTCTTGGAACTCCAGTCCCTCCAGGGCAGGACCA 420

LOCATION: 1..1059
PCT-US95-01829-1

Query Match 47.4%; Score 518.8; DB 5; Length 1062;
Best Local Similarity 99.6%; Pred. No. 1.4e-116;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGCTGACTGAATTCCTCGTGGTCAATCTCTCTAACTCAAGGCTAAGCGTG 60
|||||
Db 1 ATGAGCTGACTGAATTCCTCGTGGTCAATCTCTCTAACTCAAGGCTAAGCGTG 60
QY 61 TCCAGCCCGGCTCCTCTGCTGTGAGCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
|||||
Db 61 TCCAGCCCGGCTCCTCTGCTGTGAGCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
QY 121 CATGCTCTCACAGCAGACTGAGCCAGTGGCCAGAGTTCACCCCTTTGCCCTACACCTGTC 180
|||||
Db 121 CATGCTCTCACAGCAGACTGAGCCAGTGGCCAGAGTTCACCCCTTTGCCCTACACCTGTC 180
QY 181 CTGCTGCTGCTGTGACTTTAGCTTGGGAGAAATGAAAACCCAGATGGAGGAGACCAAG 240
|||||
Db 181 CTGCTGCTGCTGTGACTTTAGCTTGGGAGAAATGAAAACCCAGATGGAGGAGACCAAG 240
QY 241 GCACAGGACATTCCTGAGCAGTGAACCTTCTGCTGGAGGAGTGTGCGAGCAGCGGGA 300
|||||
Db 241 GCACAGGACATTCCTGAGCAGTGAACCTTCTGCTGGAGGAGTGTGCGAGCAGCGGGA 300
QY 301 CAACTGGGACCACTTGCCTCTCATCCCTCTGCTGGGCGAGTTTCTGGACAGGTCCTGTC 360
|||||
Db 301 CAACTGGGACCACTTGCCTCTCATCCCTCTGCTGGGCGAGTTTCTGGACAGGTCCTGTC 360
QY 361 CTCTTTGGGGCCCTGAGAGCCTCTTGAACCCAGCTTCCACACCTGCTCCAGGAGGACCA 420
|||||
Db 361 CTCTTTGGGGCCCTGAGAGCCTCTTGAACCCAGCTTCCACACCTGCTCCAGGAGGACCA 420
QY 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
|||||
Db 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
QY 481 CGTTTCTGATGCTTAGAGGGTCCACCCCTCTGGGTACGG 522
|||||
Db 481 CGTTTCTGATGCTTAGAGGGTCCACCCCTCTGGGTACGG 522

RESULT 10

PCT-US95-14932-3
; Sequence 3, Application PC/TUS9514932
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; APPLICANT: 1201 Eastlake Avenue East
; APPLICANT: Seattle
; APPLICANT: WA
; APPLICANT: USA
; APPLICANT: 98102
; TITLE OF INVENTION: METHOD FOR SECRETING THROMBOPOIETIN
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/14932
; FILING DATE:

CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-648
; REFERENCE/DOCKET NUMBER: 94-13PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6673
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1062 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1059
; PCT-US95-14932-3

Query Match 47.4%; Score 518.8; DB 5; Length 1062;
Best Local Similarity 99.6%; Pred. No. 1.4e-116;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGCTGACTGAATTCCTCGTGGTCAATCTCTCTAACTCAAGGCTAAGCGTG 60
|||||
Db 1 ATGAGCTGACTGAATTCCTCGTGGTCAATCTCTCTAACTCAAGGCTAAGCGTG 60
QY 61 TCCAGCCCGGCTCCTCTGCTGTGAGCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
|||||
Db 61 TCCAGCCCGGCTCCTCTGCTGTGAGCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
QY 121 CATGCTCTCACAGCAGACTGAGCCAGTGGCCAGAGTTCACCCCTTTGCCCTACACCTGTC 180
|||||
Db 121 CATGCTCTCACAGCAGACTGAGCCAGTGGCCAGAGTTCACCCCTTTGCCCTACACCTGTC 180
QY 181 CTGCTGCTGCTGTGACTTTAGCTTGGGAGAAATGAAAACCCAGATGGAGGAGACCAAG 240
|||||
Db 181 CTGCTGCTGCTGTGACTTTAGCTTGGGAGAAATGAAAACCCAGATGGAGGAGACCAAG 240
QY 241 GCACAGGACATTCCTGAGCAGTGAACCTTCTGCTGGAGGAGTGTGCGAGCAGCGGGA 300
|||||
Db 241 GCACAGGACATTCCTGAGCAGTGAACCTTCTGCTGGAGGAGTGTGCGAGCAGCGGGA 300
QY 301 CAACTGGGACCACTTGCCTCTCATCCCTCTGCTGGGCGAGTTTCTGGACAGGTCCTGTC 360
|||||
Db 301 CAACTGGGACCACTTGCCTCTCATCCCTCTGCTGGGCGAGTTTCTGGACAGGTCCTGTC 360
QY 361 CTCTTTGGGGCCCTGAGAGCCTCTTGAACCCAGCTTCCACACCTGCTCCAGGAGGACCA 420
|||||
Db 361 CTCTTTGGGGCCCTGAGAGCCTCTTGAACCCAGCTTCCACACCTGCTCCAGGAGGACCA 420
QY 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
|||||
Db 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
QY 481 CGTTTCTGATGCTTAGAGGGTCCACCCCTCTGGGTACGG 522
|||||
Db 481 CGTTTCTGATGCTTAGAGGGTCCACCCCTCTGGGTACGG 522

RESULT 11

PCT-US95-16626-3
; Sequence 3, Application PC/TUS9516626
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; APPLICANT: 1201 Eastlake Avenue East
; APPLICANT: Seattle
; APPLICANT: WA
; APPLICANT: USA
; APPLICANT: 98102
; TITLE OF INVENTION: PURIFIED THROMBOPOIETIN AND METHOD OF

;; TITLE OF INVENTION: MAKING IT
;; NUMBER OF SEQUENCES: 24
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: ZymoGenetics, Inc.
;; STREET: 1201 Eastlake Avenue East
;; CITY: Seattle
;; STATE: WA
;; COUNTRY: USA
;; ZIP: 98102
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US95/16626
;; FILING DATE:
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Parker, Gary E
;; REGISTRATION NUMBER: 31,648
;; REFERENCE/DOCKET NUMBER: 94-11PC
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 206-442-6673
;; TELEFAX: 206-442-6678
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1062 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: cdna
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 1..1059
PCT-US95-16626-3

Query Match 47.4%; Score 518.8; DB 5; Length 1062;
Best Local Similarity 99.6%; Pred. No. 1.4e-116;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGAGCTGACTGAATTGCTCTCGTGGTCATGCTTCTCTAACTGCAAGGCTAACGCTG 60
Db 1 ATGAGCTGACTGAATTGCTCTCGTGGTCATGCTTCTCTAACTGCAAGGCTAACGCTG 60
Qy 61 TCCAGCCCGGCTCTCTCTGCTTGTGACCTCCGAGTCTCAGTAACTGCTTGTGACTCC 120
Db 61 TCCAGCCCGGCTCTCTCTGCTTGTGACCTCCGAGTCTCAGTAACTGCTTGTGACTCC 120
Qy 121 CATGTCTCTCACAGCAGCTGAGCCAGTGCAGGAGTTTCAACCTTTTGGCTTACACTGTC 180
Db 121 CATGTCTCTCACAGCAGCTGAGCCAGTGCAGGAGTTTCAACCTTTTGGCTTACACTGTC 180
Qy 181 CTGCTGCTCTGTGGACTTTAGCTTTGGAGATGAAACCCAGATGGAGGAGACCAAG 240
Db 181 CTGCTGCTCTGTGGACTTTAGCTTTGGAGATGAAACCCAGATGGAGGAGACCAAG 240
Qy 241 GCACAGACATTTGGGAGCAGTACCTTCTGCTGGAGGAGTGTATGCGACACGGGA 300
Db 241 GCACAGACATTTGGGAGCAGTACCTTCTGCTGGAGGAGTGTATGCGACACGGGA 300
Qy 301 CAACCTGGAGCCACTTGCCTCTCATCCCTCTGGGAGCTTTCTGGACAGTCTGCTCTC 360
Db 301 CAACCTGGAGCCACTTGCCTCTCATCCCTCTGGGAGCTTTCTGGACAGTCTGCTCTC 360
Qy 361 CTGCTTTGGGCCCCGAGAGCTCTTTGGAACCCAGTCTTCTCCACAGGGCAGGACCA 420
Db 361 CTGCTTTGGGCCCCGAGAGCTCTTTGGAACCCAGTCTTCTCCACAGGGCAGGACCA 420
Qy 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCAACTCTGCTCCGAGGAAGGTG 480
Db 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCAACTCTGCTCCGAGGAAGGTG 480

Qy 481 CGTTTCTCTGATGCTCTTAGAGAGGTTCCACCTCTCGGTACGG 522
Db 481 CGTTTCTCTGATGCTCTTAGAGAGGTTCCACCTCTCGGTACGG 522

RESULT 12

US-08-388-779A-1
; Sequence 1, Application US/08388779A
; Patent No. 5696250
; GENERAL INFORMATION:
; APPLICANT: Elliott, Steven G.
; TITLE OF INVENTION: MGF ANALOGS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: 1840 Behavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1720
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/388,779A
; FILING DATE: 15-FEB-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Cook Ph.D., Robert R.
; REGISTRATION NUMBER: 31,602
; REFERENCE/DOCKET NUMBER: A-337
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1342 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36..1094
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 99..1094
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 36..98
US-08-388-779A-1

Query Match 47.4%; Score 518.8; DB 1; Length 1342;
Best Local Similarity 99.6%; Pred. No. 1.5e-116;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGAGCTGACTGAATTGCTCTCGTGGTCATGCTTCTCTAACTGCAAGGCTAACGCTG 60
Db 36 ATGGAGCTGACTGAATTGCTCTCGTGGTCATGCTTCTCTAACTGCAAGGCTAACGCTG 95
Qy 61 TCCAGCCCGGCTCTCTCTGCTTGTGACCTCCGAGTCTCAGTAACTGCTTGTGACTCC 120
Db 96 TCCAGCCCGGCTCTCTCTGCTTGTGACCTCCGAGTCTCAGTAACTGCTTGTGACTCC 155
Qy 121 CATGTCTCTCACAGCAGCTGAGCCAGTGCAGGAGTTTCAACCTTTGCTTACACTGTC 180
Db 156 CATGTCTCTCACAGCAGCTGAGCCAGTGCAGGAGTTTCAACCTTTGCTTACACTGTC 215
Qy 181 CTGCTGCTCTGTGCTGCTTGTGAGCTTTAGCTTTGGAGATGAAACCCAGATGGAGGAGACCAAG 240
Db 216 CTGCTGCTCTGTGCTGCTTGTGAGCTTTAGCTTTGGAGATGAAACCCAGATGGAGGAGACCAAG 275


```
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/252,628
;; FILING DATE: 31-MAY-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/321,488
;; FILING DATE: 12-OCT-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/347,780
;; FILING DATE: 30-NOV-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Cook Ph.D., Robert R.
;; REGISTRATION NUMBER: 31,602
;; REFERENCE/DOCKET NUMBER: A-290D
;; INFORMATION FOR SEQ ID NO: 24:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1342 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: unknown
;; TOPOLOGY: unknown
;; MOLECULE TYPE: CDNA
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 36..1097
;; FEATURE:
;; NAME/KEY: mat_peptide
;; LOCATION: 99..1097
;; FEATURE:
;; NAME/KEY: sig_peptide
;; LOCATION: 36..98
;;
US-08-413-803-24
```

```
Query Match 47.4%; Score 518.8; DB 1; Length 1342;
Best Local Similarity 99.6%; Pred No. 1.5e-116;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATTCCTCGTGGTCAATGCTTCTCTAACTGCAAGGCTAACGCTG 60
Db 36 ATGGAGCTGACTGAATTCCTCGTGGTCAATGCTTCTCTAACTGCAAGGCTAACGCTG 95
QY 61 TCCAGCCCGGGCTCCCTCGTGGTCAATGCTTCTCTAACTGCAAGGCTAACGCTG 120
Db 96 TCCAGCCCGGGCTCCCTCGTGGTCAATGCTTCTCTAACTGCAAGGCTAACGCTG 155
QY 121 CATGTCCTTCACAGAGACTGAGCCAGTCCGAGAGTTCACCTTTGCTACACCTGTC 180
Db 156 CATGTCCTTCACAGAGACTGAGCCAGTCCGAGAGTTCACCTTTGCTACACCTGTC 215
QY 181 CTGTCCTGCTGTGGACTTTAGCTTGGGAGATGGAATGGAATGGAATGGAATGGAATG 240
Db 216 CTGTCCTGCTGTGGACTTTAGCTTGGGAGATGGAATGGAATGGAATGGAATGGAATG 275
QY 241 GCACAGGACATTCGGGAGCAGTGACCTTCTGCTGGAGGAGTGTGTCAGCAGCGGGA 300
Db 276 GCACAGGACATTCGGGAGCAGTGACCTTCTGCTGGAGGAGTGTGTCAGCAGCGGGA 335
QY 301 CAACTGGGACCCACTTGCCTCTCATCCCTCTGCTGGGAGAGCTTCTGAGACAGTCCGTC 360
Db 336 CAACTGGGACCCACTTGCCTCTCATCCCTCTGCTGGGAGAGCTTCTGAGACAGTCCGTC 395
QY 361 CTCCTTGGGGCCCTGCAGAGCCTTCTTGAACCCAGCTTCTTCCACAGGGGAGGACACA 420
Db 396 CTCCTTGGGGCCCTGCAGAGCCTTCTTGAACCCAGCTTCTTCCACAGGGGAGGACACA 455
QY 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCAACTGCTCCGAGGAAAGGTG 480
Db 456 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCAACTGCTCCGAGGAAAGGTG 515
QY 481 CGTTTCTGATGCTGTAGGAGGCTCCACCTCTCGGTACGG 522
Db 516 CGTTTCTGATGCTGTAGGAGGCTCCACCTCTCGGTACGG 557
```

```
RESULT 15
US-08-321-488A-24
; Sequence 24, Application US/08321488A
; Patent No. 5795569
; GENERAL INFORMATION:
; APPLICANT: Bartley, Timothy D.
; APPLICANT: Bogenberger, Jakob M.
; APPLICANT: Bosselman, Robert A.
; APPLICANT: Hunt, Pamela
; APPLICANT: Kinstler, Olaf B.
; APPLICANT: Samal, Babru B.
; TITLE OF INVENTION: MONO-PEGYLATED PROTEINS THAT STIMULATE
; TITLE OF INVENTION: MEGAKARYOCYTE GROWTH AND DIFFERENTIATION
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: US
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/321,488A
; FILING DATE: 12-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/252,628
; FILING DATE: 31-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,768
; FILING DATE: 31-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Cook, Robert R.
; REGISTRATION NUMBER: 31,602
; REFERENCE/DOCKET NUMBER: A-290B
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1342 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36..1097
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 99..1097
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 36..98
;
US-08-321-488A-24

Query Match 47.4%; Score 518.8; DB 1; Length 1342;
Best Local Similarity 99.6%; Pred No. 1.5e-116;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

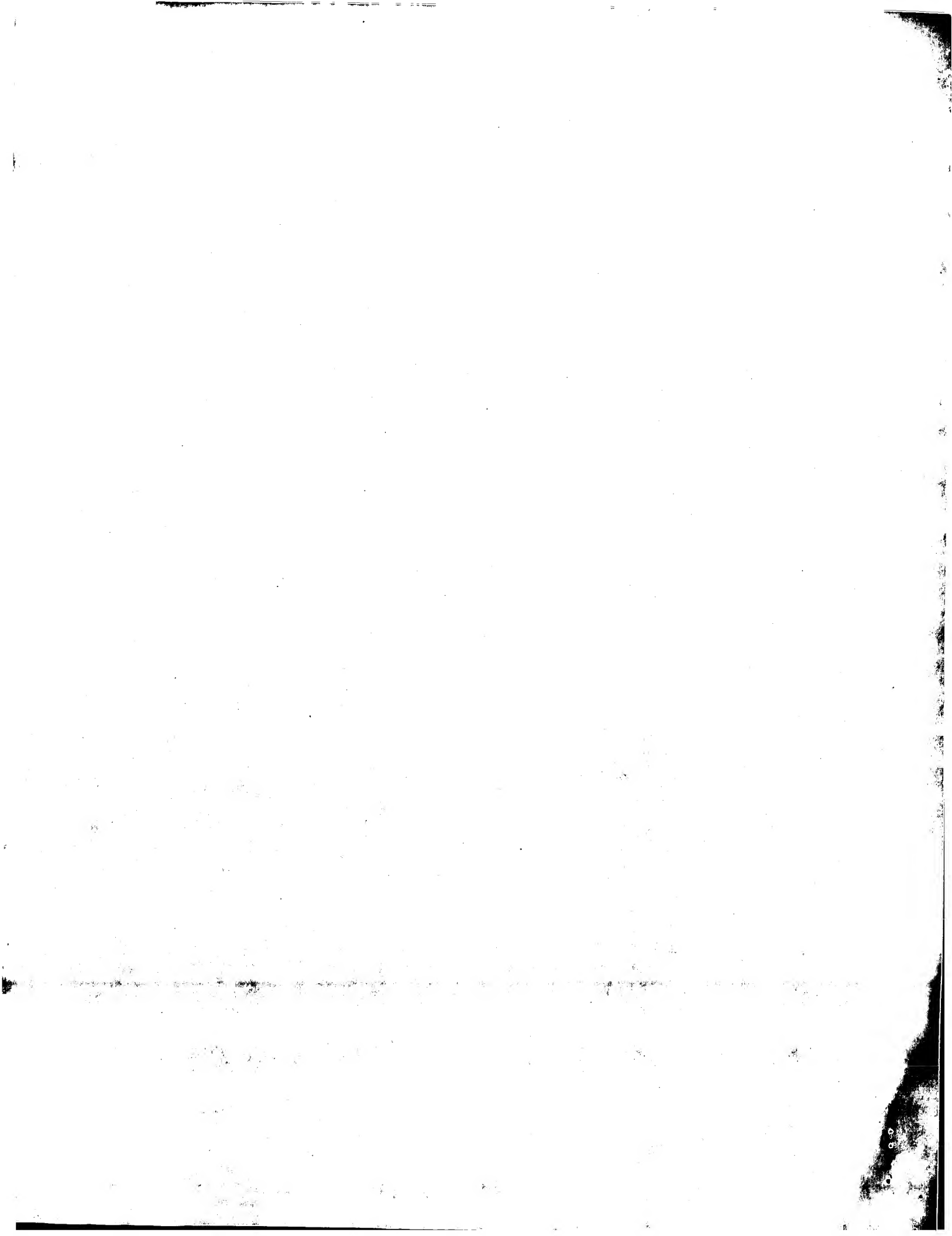
QY 1 ATGGAGCTGACTGAATTCCTCGTGGTCAATGCTTCTCTAACTGCAAGGCTAACGCTG 60
Db 36 ATGGAGCTGACTGAATTCCTCGTGGTCAATGCTTCTCTAACTGCAAGGCTAACGCTG 95
QY 61 TCCAGCCCGGGCTCCCTCGTGGTCAATGCTTCTCTAACTGCAAGGCTAACGCTG 120
Db 96 TCCAGCCCGGGCTCCCTCGTGGTCAATGCTTCTCTAACTGCAAGGCTAACGCTG 155
QY 121 CATGTCCTTCACAGAGACTGAGCCAGTCCGAGAGTTCACCTTTGCTACACCTGTC 180
```

```

Db 156 CATGTCCTTACAGCAGACTGAGCCAGTGCCAGAGTTACCCCTTTCCTACACCTGTC 215
Qy 181 CTGCTGCTGCTGTGGACTTTAGCTTGGAGAAATGGAACCCACAGATGGAGGAGACCAAG 240
Db 216 CTGCTGCTGCTGTGGACTTTAGCTTGGAGAAATGGAACCCACAGATGGAGGAGACCAAG 275
Qy 241 GCACAGGACATTCCTGGGAGCAGTGACCCCTTCTGTGGAGGAGTGATGGCAGCAGCGGGA 300
Db 276 GCACAGGACATTCCTGGGAGCAGTGACCCCTTCTGTGGAGGAGTGATGGCAGCAGCGGGA 335
Qy 301 CAACTGGGACCCACTTGCCTCTCATCCCTCTGCTGGGCGAGCTTCTGGGACAGGTCGCTC 360
Db 336 CAACTGGGACCCACTTGCCTCTCATCCCTCTGCTGGGCGAGCTTCTGGGACAGGTCGCTC 395
Qy 361 CTCCTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGCGAGGACCA 420
Db 396 CTCCTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGCGAGGACCA 455
Qy 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGGTG 480
Db 456 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGGTG 515
Qy 481 CGTTTCCTGATGCTGTAGGAGGTTCCACCCCTCTCGGTACGG 522
Db 516 CGTTTCCTGATGCTGTAGGAGGTTCCACCCCTCTCGGTACGG 557

```

Search completed: April 5, 2002, 10:35:04
 Job time: 4998 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2002, 10:34:53 ; Search time 81.22 Seconds
(without alignments)
3019.893 Million cell updates/sec

Title: US-09-680-514-6

Perfect score: 1083

Sequence: 1 ATGGAGCTGACTGAATTGCT.....TACGCCACTTCCAGCC 1083

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	523.2	48.3	1059	US-08-471-045-57	Sequence 57, Appl
2	523.2	48.3	1059	US-08-469-712A-57	Sequence 57, Appl
3	523.2	48.3	1062	US-08-347-029-3	Sequence 3, Appl
4	523.2	48.3	1062	US-08-457-254-3	Sequence 3, Appl
5	523.2	48.3	1062	US-08-484-257-18	Sequence 18, Appl
6	523.2	48.3	1062	US-08-482-212-1	Sequence 1, Appl
7	523.2	48.3	1062	US-08-999-927-1	Sequence 1, Appl
8	523.2	48.3	1062	PCT-US94-08806-18	Sequence 18, Appl
9	523.2	48.3	1062	PCT-US95-01829-1	Sequence 1, Appl
10	523.2	48.3	1062	PCT-US95-14932-3	Sequence 3, Appl
11	523.2	48.3	1062	PCT-US95-16626-3	Sequence 3, Appl
12	523.2	48.3	1342	US-08-388-779A-1	Sequence 1, Appl
13	523.2	48.3	1342	US-08-591-070A-1	Sequence 1, Appl
14	523.2	48.3	1342	US-08-413-803-24	Sequence 24, Appl
15	523.2	48.3	1342	US-08-321-488A-24	Sequence 24, Appl
16	523.2	48.3	1342	US-08-414-161B-1	Sequence 1, Appl
17	523.2	48.3	1342	US-08-927-855-1	Sequence 1, Appl
18	523.2	48.3	1342	PCT-US95-03776-24	Sequence 24, Appl
19	523.2	48.3	1342	PCT-US95-03776-26	Sequence 26, Appl
20	523.2	48.3	1795	US-08-330-517-1	Sequence 1, Appl
21	521.6	48.2	605	US-08-388-779A-3	Sequence 3, Appl
22	521.6	48.2	605	US-08-591-070A-3	Sequence 3, Appl
23	521.6	48.2	605	US-08-927-855-3	Sequence 3, Appl
24	486.8	44.9	525	US-08-434-411-1	Sequence 1, Appl
25	486.8	44.9	525	US-08-434-402-1	Sequence 1, Appl
26	486.8	44.9	525	US-08-783-288-1	Sequence 1, Appl
27	486.8	44.9	525	US-08-890-640-1	Sequence 1, Appl

28 486.8 44.9 525 6 5194592-25 Patent No. 5194592
29 486.6 44.9 1017 3 US-08-469-318-67 Sequence 67, Appl
30 486.6 44.9 1017 3 US-08-468-609A-67 Sequence 67, Appl
31 486.6 44.9 1017 5 PCT-US95-01185-67 Sequence 67, Appl
32 478 44.1 1164 1 US-08-413-803-26 Sequence 26, Appl
33 478 44.1 1164 1 US-08-321-488A-26 Sequence 26, Appl
34 478 44.1 1164 5 PCT-US95-03776-28 Sequence 28, Appl
35 477.6 44.1 921 3 US-08-469-318-84 Sequence 84, Appl
36 477.6 44.1 921 3 US-08-468-609A-84 Sequence 84, Appl
37 477.6 44.1 921 3 PCT-US95-01185-84 Sequence 84, Appl
38 476.2 44.0 546 3 US-08-489-318-177 Sequence 177, Appl
39 476.2 44.0 546 3 US-08-469-318-178 Sequence 178, Appl
40 476.2 44.0 546 3 US-08-468-609A-177 Sequence 177, Appl
41 476.2 44.0 546 3 US-08-468-609A-178 Sequence 178, Appl
42 476.2 44.0 546 5 PCT-US95-01185-177 Sequence 177, Appl
43 476.2 44.0 546 5 PCT-US95-01185-178 Sequence 178, Appl
44 476.2 44.0 921 3 US-08-469-318-72 Sequence 72, Appl
45 476.2 44.0 921 3 US-08-489-318-75 Sequence 75, Appl

ALIGNMENTS

RESULT 1

US-08-471-045-57
; Sequence 57, Application US/08471045
; Patent No. 6060047
; GENERAL INFORMATION:
; APPLICANT: Abrams, Mark A.
; APPLICANT: Bauer, S. C.
; APPLICANT: Braford-Goldberg, Sarah R.
; APPLICANT: Caparon, Mairé H.
; APPLICANT: Easton, Alan M.
; APPLICANT: Klein, Barbara K.
; APPLICANT: McKearn, John P.
; APPLICANT: Olin, Peter O.
; APPLICANT: Paik, Kumnan
; APPLICANT: Thomas, John W.
; TITLE OF INVENTION: Co-administration of Interleukin-3
; TITLE OF INVENTION: Mutant Polypeptides with CSF's for Multi-lineage
; TITLE OF INVENTION: Hematopoietic Cell Production
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,
; ADDRESSEE: Corporate Patent Dept.
; STREET: P. O. Box 5110
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60680
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,045
; FILING DATE: 06-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/01184
; FILING DATE: 02-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/193,373
; FILING DATE: 04-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Dennis A.
; REGISTRATION NUMBER: 34,547
; REFERENCE/DOCKET NUMBER: C-2789/3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (708)470-6501
; TELEFAX: (708)470-6881
; INFORMATION FOR SEQ ID NO: 57:


```

QY 61 TCAGAGCCGGCTCCTCTGCTGTCACCTCCGAGTCCCTCACTAACTGCTTCGTGACTCC 120
DB 61 TCAGAGCCGGCTCCTCTGCTGTCACCTCCGAGTCCCTCACTAACTGCTTCGTGACTCC 120
QY 121 CATGCTCTTCACAGAGACTGAGCAGTCCGAGAGTTCACCTTCCTGCTTACACCTGTC 180
DB 121 CATGCTCTTCACAGAGACTGAGCAGTCCGAGAGTTCACCTTCCTGCTTACACCTGTC 180
QY 181 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
DB 181 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 241 GCACAGGACATTCCTGGAGCAGTACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
DB 241 GCACAGGACATTCCTGGAGCAGTACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 301 CAACCTGGAGCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
DB 301 CAACCTGGAGCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 361 CTCCTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGCGCAGGACCA 420
DB 361 CTCCTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGCGCAGGACCA 420
QY 421 GCTCACAAGGATCCCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 421 GCTCACAAGGATCCCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 481 CGTTTCCTGATGCTGTAGGAGGCTGCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 528
DB 481 CGTTTCCTGATGCTGTAGGAGGCTGCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 528

```

RESULT 5

```

US-08-484-257-18
; Sequence 18, Application US/08484257
; Patent No. 5989537
; GENERAL INFORMATION:
; APPLICANT: Holly, Richard D.
; APPLICANT: Lok, Si
; APPLICANT: Foster, Donald C.
; APPLICANT: Hagen, Frederick S.
; APPLICANT: Kaushansky, Kenneth
; APPLICANT: Kuiper, Joseph L.
; APPLICANT: Lofton-Day, Catherine E.
; APPLICANT: Oort, Pieter J.
; TITLE OF INVENTION: Methods for Stimulating Granulocyte/Macrophage Lineage Cells
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,257
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-648
; REFERENCE/DOCKET NUMBER: 95-09
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6673
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 18:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1062 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1059
; US-08-484-257-18

```

```

Query Match 48.3%; Score 523.2; DB 2; Length 1062;
Best Local Similarity 99.4%; Pred. No. 3.1e-118;
Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 ATGGAGTGACTGAATTTGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
DB 1 ATGGAGTGACTGAATTTGCTTCCTCGTGGTTCATGCTTCCTTAAGTCAAGGCTAACGCTG 60
QY 61 TCAGAGCCGGCTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
DB 61 TCAGAGCCGGCTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 121 CATGCTCTTCACAGCAGACTGAGCAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
DB 121 CATGCTCTTCACAGCAGACTGAGCAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 181 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
DB 181 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 241 GCACAGGACATTCCTGGAGCAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
DB 241 GCACAGGACATTCCTGGAGCAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 301 CAACCTGGAGCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
DB 301 CAACCTGGAGCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 361 CTCCTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGCGCAGGACCA 420
DB 361 CTCCTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGCGCAGGACCA 420
QY 421 GCTCACAAGGATCCCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 421 GCTCACAAGGATCCCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 481 CGTTTCCTGATGCTGTAGGAGGCTGCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 528
DB 481 CGTTTCCTGATGCTGTAGGAGGCTGCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 528

```

RESULT 6

```

US-08-482-212-1
; Sequence 1, Application US/08482212
; Patent No. 6013067
; GENERAL INFORMATION:
; APPLICANT: Fibbe, Willem E.
; APPLICANT: Grossmann, Angelika
; TITLE OF INVENTION: METHODS FOR INCREASING HEMATOPOIETIC CELLS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

RESULT      7
US-08-999-927-1
: Sequence 1, Application US/08999927
: Patent No. 6099830
: GENERAL INFORMATION:
: APPLICANT: Kaushansky, Kenneth
: TITLE OF INVENTION: Methods of Stimulating Erythropoiesis

```



```
; LOCATION: 1..1059
PCT-US95-01829-1

Query Match      48.3%; Score 523.2; DB 5; Length 1062;
Best Local Similarity 99.4%; Pred. No. 3.1e-118;
Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGAGCTGACTGAATGCTCTCGTGGTCAATGCTTCTCCTAACTGCAAGGCTAACGCTG 60
Db 1 ATGAGCTGACTGAATGCTCTCGTGGTCAATGCTTCTCCTAACTGCAAGGCTAACGCTG 60
QY 61 TCCAGCCGGCTCCTCGTGGTCAATGCTTGGAGTCCCTCACTAACTGCTTCGCTGACTCC 120
Db 61 TCCAGCCGGCTCCTCGTGGTCAATGCTTGGAGTCCCTCACTAACTGCTTCGCTGACTCC 120
QY 121 CATGCTCTTACAGCAGACTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 180
Db 121 CATGCTCTTACAGCAGACTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 180
QY 181 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db 181 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 241 GCACAGGACATTTCTGGAGCAGTGCACCTTCTGCTGGAGGAGTGCATGGCAGCAGGGA 300
Db 241 GCACAGGACATTTCTGGAGCAGTGCACCTTCTGCTGGAGGAGTGCATGGCAGCAGGGA 300
QY 301 CAACCTGGACCCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 301 CAACCTGGACCCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 361 CTCCTTGGGGCCCTGCGAGAGCTCTTGGAGACCCAGCTTCTCCACAGGGCAGGACCA 420
Db 361 CTCCTTGGGGCCCTGCGAGAGCTCTTGGAGACCCAGCTTCTCCACAGGGCAGGACCA 420
QY 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGTG 480
Db 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGTG 480
QY 481 CGTTTCTGATGCTTGTAGGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGGCTGAGG 528
Db 481 CGTTTCTGATGCTTGTAGGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGGCTGAGG 528

RESULT 11
PCT-US95-16626-3
; Sequence 3, Application PC/TUS9516626
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; APPLICANT: 1201 Eastlake Avenue East
; APPLICANT: Seattle
; APPLICANT: WA
; APPLICANT: USA
; APPLICANT: 98102
; TITLE OF INVENTION: PURIFIED THROMBOPOIETIN AND METHOD OF
```

;; TITLE OF INVENTION: MAKING IT
;; NUMBER OF SEQUENCES: 24
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: ZymoGenetics, Inc.
;; STREET: 1201 Eastlake Avenue East
;; CITY: Seattle
;; STATE: WA
;; COUNTRY: USA
;; ZIP: 98102
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US95/16626
;; FILING DATE:
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Parker, Gary E
;; REGISTRATION NUMBER: 31,648
;; REFERENCE/DOCKET NUMBER: 94-11PC
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 206-442-6673
;; TELEFAX: 206-442-6678
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1062 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 1..1059
;; PCT-US95-16626-3

Query Match 48.3%; Score 523.2; DB 5; Length 1062;
Best Local Similarity 99.4%; Pred. No. 3.1e-118;
Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATGCTCTCGTGTGCTATGCTTCTCTAACTCAAGGCTAAGCGTG 60
Db 1 ATGGAGCTGACTGAATGCTCTCGTGTGCTATGCTTCTCTAACTCAAGGCTAAGCGTG 60
QY 61 TCCAGCCCGGCTCTCTGCTGTGCTCGGAGTCTCAGTAACTGCTTCTGCTACTCC 120
Db 61 TCCAGCCCGGCTCTCTGCTGTGCTCGGAGTCTCAGTAACTGCTTCTGCTACTCC 120
QY 121 CATGTCTCTTACACGACACTGAGCCAGTCCAGAGGTTTACCCCTTGGCTACACCTGTC 180
Db 121 CATGTCTCTTACACGACACTGAGCCAGTCCAGAGGTTTACCCCTTGGCTACACCTGTC 180
QY 181 CTGTGCTGCTGTGGAGTCTTACCTTGGGAGATGAAACCCAGATGGAGGACCAAG 240
Db 181 CTGTGCTGCTGTGGAGTCTTACCTTGGGAGATGAAACCCAGATGGAGGACCAAG 240
QY 241 GCACAGACATTTGGGAGCAGTGACCTTCTGCTGGAGGAGTATGGCAGCAGCGGGA 300
Db 241 GCACAGACATTTGGGAGCAGTGACCTTCTGCTGGAGGAGTATGGCAGCAGCGGGA 300
QY 301 CAACCTGGGACCCACTTGCCTTCTATCCCTCGCTGGGAGCTTCTGGAGAGTCCGCTCTC 360
Db 301 CAACCTGGGACCCACTTGCCTTCTATCCCTCGCTGGGAGCTTCTGGAGAGTCCGCTCTC 360
QY 361 CTCTTGGGCGCTCGAGAGCCCTTGGAAACCCAGCTTCTTCCACAGGGCAGGACCA 420
Db 361 CTCTTGGGCGCTCGAGAGCCCTTGGAAACCCAGCTTCTTCCACAGGGCAGGACCA 420
QY 421 GCTCACAAGGATCCCAATGCCATCTCTCTGAGCTTCCAAACACTCTGCTCCGAGGAAGTG 480
Db 421 GCTCACAAGGATCCCAATGCCATCTCTCTGAGCTTCCAAACACTCTGCTCCGAGGAAGTG 480

QY 481 CGTTTCCTGATGCTGTAGGAGGTCCACCTCTCGTCAGGGGTGGC 528
Db 481 CGTTTCCTGATGCTGTAGGAGGTCCACCTCTCGTCAGGGGTGGC 528
RESULT 12
US-08-388-779A-1
;; Sequence 1, Application US/08388779A
;; Patent No. 5696250
;; GENERAL INFORMATION:
;; APPLICANT: Elliott, Steven G.
;; TITLE OF INVENTION: MGDF ANALOGS
;; NUMBER OF SEQUENCES: 17
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: AMGEN INC.
;; STREET: 1840 Dehavilland Drive
;; CITY: Thousand Oaks
;; STATE: California
;; COUNTRY: USA
;; ZIP: 91320-1720
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/388,779A
;; FILING DATE: 15-FEB-1995
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Cook Ph.D., Robert R.
;; REGISTRATION NUMBER: 31,602
;; REFERENCE/DOCKET NUMBER: A-337
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1342 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 36..1094
;; FEATURE:
;; NAME/KEY: mat_peptide
;; LOCATION: 99..1094
;; FEATURE:
;; NAME/KEY: sig_peptide
;; LOCATION: 36..98
;; US-08-388-779A-1

Query Match 48.3%; Score 523.2; DB 1; Length 1342;
Best Local Similarity 99.4%; Pred. No. 3.3e-118;
Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATGCTCTCGTGTGCTATGCTTCTCTAACTCAAGGCTAAGCGTG 60
Db 36 ATGGAGCTGACTGAATGCTCTCGTGTGCTATGCTTCTCTAACTCAAGGCTAAGCGTG 95
QY 61 TCCAGCCCGGCTCTCTGCTGTGCTCGGAGTCTCAGTAACTGCTTCTGCTACTCC 120
Db 96 TCCAGCCCGGCTCTCTGCTGTGCTCGGAGTCTCAGTAACTGCTTCTGCTACTCC 155
QY 121 CATGTCTCTTACACGACACTGAGCCAGTCCAGAGGTTTACCCCTTGGCTACACCTGTC 180
Db 156 CATGTCTCTTACACGACACTGAGCCAGTCCAGAGGTTTACCCCTTGGCTACACCTGTC 215
QY 181 CTGTGCTGCTGTGGAGTCTTACCTTGGGAGATGAAACCCAGATGGAGGACCAAG 240
Db 216 CTGTGCTGCTGTGGAGTCTTACCTTGGGAGATGAAACCCAGATGGAGGACCAAG 275

QY 241 GCACAGGACATTCGGGAGCAGTACACCTTCTGCTGGAGGAGTGATGGCAGCAGCGGGA 300
Db 276 GCACAGGACATTCGGGAGCAGTACACCTTCTGCTGGAGGAGTGATGGCAGCAGCGGGA 335
QY 301 CAACCTGGGACCCACTTGGCTCTCATCCCTCTGCTGGGAGAGCTTCTGGACAGGTCCGTCTC 360
Db 336 CAACCTGGGACCCACTTGGCTCTCATCCCTCTGCTGGGAGAGCTTCTGGACAGGTCCGTCTC 395
QY 361 CTCCTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCACA 420
Db 396 CTCCTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCACA 455
QY 421 GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGTG 480
Db 456 GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGTG 515
QY 481 CGTTTCTGTGCTGTAGGAGGCTCCACCTCTGCTCAGGGGTGGC 528
Db 516 CGTTTCTGTGCTGTAGGAGGCTCCACCTCTGCTCAGGGGTGGC 563

RESULT 13

US-08-591-070A-1
; Sequence 1, Application US/08591070A
; Patent No. 5756083
; GENERAL INFORMATION:
; APPLICANT: Elliott, Steven G.
; TITLE OF INVENTION: MPL LIGAND ANALOGS
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: CA
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,070A
; FILING DATE: 09-FEB-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/388,779
; FILING DATE: 15-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: COOK, ROBERT R
; REGISTRATION NUMBER: 31,602
; REFERENCE/DOCKET NUMBER: A-337B
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1342 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36..1094
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 99..1094
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 36..98
US-08-591-070A-1

Query Match 48.3%; Score 523.2; DB 1; Length 1342;
Best Local Similarity 99.4%; Pred. No. 3.3e-118;

Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATGAGCTGACTGAATTCGCTCCTCGTGTGCTCATGCTTCTCCTAACTGCAAGGCTAACGCTG 60
Db 36 ATGAGCTGACTGAATTCGCTCCTCGTGTGCTCATGCTTCTCCTAACTGCAAGGCTAACGCTG 95
QY 61 TCCAGCCCGGCTCCTCCTGCTTGTGACCTCCGAGTCCCTCAGTAAACTGCTTCGTGACTCC 120
Db 96 TCCAGCCCGGCTCCTCCTGCTTGTGACCTCCGAGTCCCTCAGTAAACTGCTTCGTGACTCC 155
QY 121 CATGCTCTTACACAGACAGTACGAGTCCGAGGTTTACCCCTTTGGCTTACACCTGTC 180
Db 156 CATGCTCTTACACAGACAGTACGAGTCCGAGGTTTACCCCTTTGGCTTACACCTGTC 215
QY 181 CTGCTGCTGCTGTGGAGCTTTAGCTTGGGAGAAATGGAAACCCAGATGGAGGAGACCAAG 240
Db 216 CTGCTGCTGCTGTGGAGCTTTAGCTTGGGAGAAATGGAAACCCAGATGGAGGAGACCAAG 275
QY 241 GCACAGGACATTCCTGGAGCAGTACCCCTTCTGCTGGAGGAGTGATGGCAGCAGCGGGA 300
Db 276 GCACAGGACATTCCTGGAGCAGTACCCCTTCTGCTGGAGGAGTGATGGCAGCAGCGGGA 335
QY 301 CAACTGGGACCCACTTGGCTCTCATCCCTCTGCTGGGAGAGCTTCTGGACAGGTCCGTCTC 360
Db 336 CAACTGGGACCCACTTGGCTCTCATCCCTCTGCTGGGAGAGCTTCTGGACAGGTCCGTCTC 395
QY 361 CTCCTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCACA 420
Db 396 CTCCTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCACA 455
QY 421 GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGTG 480
Db 456 GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGTG 515
QY 481 CGTTTCTGTGCTGTAGGAGGCTCCACCTCTGCTCAGGGGTGGC 528
Db 516 CGTTTCTGTGCTGTAGGAGGCTCCACCTCTGCTCAGGGGTGGC 563

RESULT 14

US-08-413-803-24
; Sequence 24, Application US/08413803
; Patent No. 5766581
; GENERAL INFORMATION:
; APPLICANT: Bartley, Timothy D.
; APPLICANT: Bogenberger, Jakob M.
; APPLICANT: Bosselman, Robert A.
; APPLICANT: Hunt, Pamela
; APPLICANT: Kinstler, Olaf B.
; APPLICANT: Samal, Babru B.
; TITLE OF INVENTION: METHODS FOR TREATING MAMMALS WITH
; TITLE OF INVENTION: MONO-PGYLATED PROTEINS THAT STIMULATE MEGAKARYOCYTE
; TITLE OF INVENTION: GROWTH AND DIFFERENTIATION
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: US
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/413,803
; FILING DATE: 30-MAR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,768
; FILING DATE: 31-MAR-1994

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/252,628
; FILING DATE: 31-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/321,488
; FILING DATE: 12-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/347,780
; FILING DATE: 30-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Cook Ph.D., Robert R.
; REGISTRATION NUMBER: 31,602
; REFERENCE/DOCKET NUMBER: A-290D
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1342 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36..1097
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 99..1097
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 36..98
;
US-08-413-803-24

```

```

Query Match      48.3%; Score 523.2; DB 1; Length 1342;
Best Local Similarity 99.4%; Pred. No. 3.3e-118;
Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGGAGCTGACTGAATTCCTCGTGGTGCATGCTTCTCTAACTCAAGGCTAACGCTG 60
Db 36 ATGGAGCTGACTGAATTCCTCGTGGTGCATGCTTCTCTAACTCAAGGCTAACGCTG 95
Qy 61 TCCAGCCCGGCTCCTCTGCTGACCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
Db 96 TCCAGCCCGGCTCCTCTGCTGACCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 155
Qy 121 CATGCTGCTGCTGGAGCTGAGCCAGTGCAGAGAGTTTCCAGGCTTTCGTACACCTGTC 180
Db 156 CATGCTGCTGCTGGAGCTGAGCCAGTGCAGAGAGTTTCCAGGCTTTCGTACACCTGTC 215
Qy 181 CTGCTGCTGCTGGAGCTGAGCCAGTGCAGAGAGTTTCCAGGCTTTCGTACACCTGTC 240
Db 216 CTGCTGCTGCTGGAGCTGAGCCAGTGCAGAGAGTTTCCAGGCTTTCGTACACCTGTC 275
Qy 241 GCACAGGACATTCGGAGCAGTGCAGGCTTCTGCTGGAGGAGTGATGGCAGCAGCGGGA 300
Db 276 GCACAGGACATTCGGAGCAGTGCAGGCTTCTGCTGGAGGAGTGATGGCAGCAGCGGGA 335
Qy 301 CAACCTGGGACCCACTTGCCTCTCATCCTCTGCGGCGAGCTTTCGTGACAGGTCCTCTC 360
Db 336 CAACCTGGGACCCACTTGCCTCTCATCCTCTGCGGCGAGCTTTCGTGACAGGTCCTCTC 395
Qy 361 CTCCTTGGGGCCCTGCGAGAGCTCCTTGGAAACCCAGCTTCCCTCCACAGGCGCAGGACACA 420
Db 396 CTCCTTGGGGCCCTGCGAGAGCTCCTTGGAAACCCAGCTTCCCTCCACAGGCGCAGGACACA 455
Qy 421 GCTCACAGGATCCCATGCGATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
Db 456 GCTCACAGGATCCCATGCGATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 515
Qy 481 CGTTTCCTGATGCTGTAGGAGGCTCCACCTCTCTGCTGCTGAGGCTGGC 528
Db 516 CGTTTCCTGATGCTGTAGGAGGCTCCACCTCTCTGCTGCTGAGGCTGGC 563

```

```

RESULT 15
US-08-321-488A-24
; Sequence 24, Application US/08321488A
; Patent No. 5795569
; GENERAL INFORMATION:
; APPLICANT: Bartley, Timothy D.
; APPLICANT: Bogenberger, Jakob M.
; APPLICANT: Bosselman, Robert A.
; APPLICANT: Hunt, Pamela
; APPLICANT: Kinstler, Olaf B.
; APPLICANT: Samal, Babru B.
; TITLE OF INVENTION: MONO-PEGYLATED PROTEINS THAT STIMULATE
; TITLE OF INVENTION: MEGAKARYOCYTE GROWTH AND DIFFERENTIATION
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: US
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/321,488A
; FILING DATE: 12-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/252,628
; FILING DATE: 31-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,768
; FILING DATE: 31-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Cook, Robert R.
; REGISTRATION NUMBER: 31,602
; REFERENCE/DOCKET NUMBER: A-290B
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1342 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36..1097
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 99..1097
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 36..98
;
US-08-321-488A-24

```

```

Query Match      48.3%; Score 523.2; DB 1; Length 1342;
Best Local Similarity 99.4%; Pred. No. 3.3e-118;
Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGGAGCTGACTGAATTCCTCGTGGTGCATGCTTCTCTAACTCAAGGCTAACGCTG 60
Db 36 ATGGAGCTGACTGAATTCCTCGTGGTGCATGCTTCTCTAACTCAAGGCTAACGCTG 95
Qy 61 TCCAGCCCGGCTCCTCTGCTGACCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
Db 96 TCCAGCCCGGCTCCTCTGCTGACCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 155
Qy 121 CATGCTGCTTTCACAGCAGACTGAGCCAGTGCAGAGGTTTCCACCTTTTGCCTTACACCTGTC 180

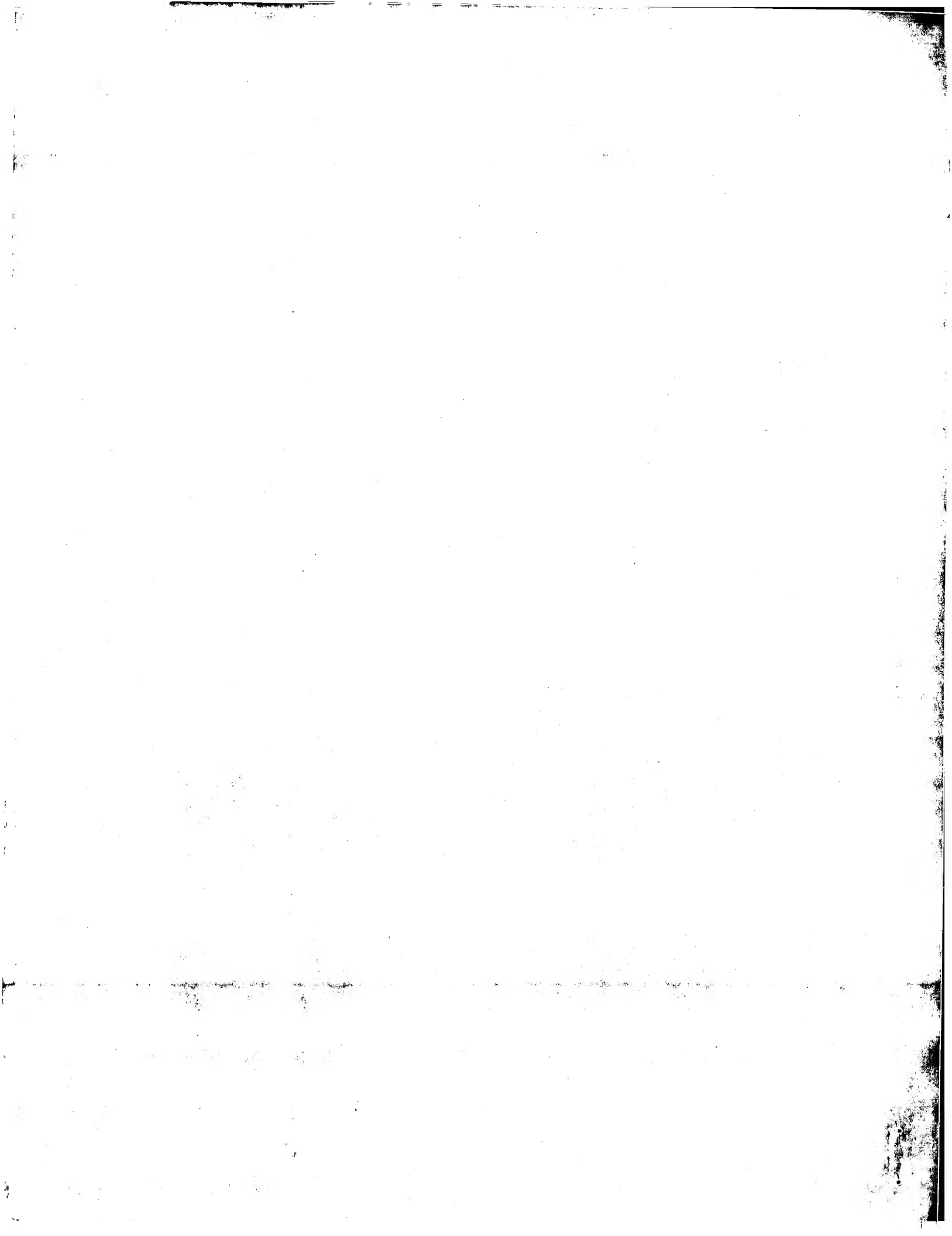
```

```

Db 156 CATGTCCCTTCACAGCAGACTGAGCCAGTGGCCAGAGGTTACCCCTTTCCTACACCTGTC 215
Qy 181 CTGCTGCCCTGCTGTGGACTTTAGCTTGGAGAAATGGAACCCAGATGGAGGAGACCAAG 240
Db 216 CTGCTGCCCTGCTGTGGACTTTAGCTTGGAGAAATGGAACCCAGATGGAGGAGACCAAG 275
Qy 241 GCACAGGACATCTGGGAGCAGTACCCTTCTGTGGAGGAGTGTATGCCAGCAGCGGGGA 300
Db 276 GCACAGGACATCTGGGAGCAGTACCCTTCTGTGGAGGAGTGTATGCCAGCAGCGGGGA 335
Qy 301 CAACCTGGGACCCACTTGGCTCTCATCCCTCTGGGGCAGCTTCTGGACAGGTCCGCTC 360
Db 336 CAACCTGGGACCCACTTGGCTCTCATCCCTCTGGGGCAGCTTCTGGACAGGTCCGCTC 395
Qy 361 CTCCTTGGGGCCCTGCAGAGCTCTCTTGGAAACCAGCTTCTCCACAGGGCAGGACCACA 420
Db 396 CTCCTTGGGGCCCTGCAGAGCTCTCTTGGAAACCAGCTTCTCCACAGGGCAGGACCACA 455
Qy 421 GCTCACAAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACCTGCTCCGAGGAAAGGTG 480
Db 456 GCTCACAAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACCTGCTCCGAGGAAAGGTG 515
Qy 481 CGTTTCCTGATGCTGTAGGAGGGTCCACCCTCTGCGTCAGGGGTGGC 528
Db 516 CGTTTCCTGATGCTGTAGGAGGGTCCACCCTCTGCGTCAGGGGTGGC 563

```

Search completed: April 5, 2002, 10:34:58
 Job time: 4992 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2002, 09:11:46 ; Search time 81.22 Seconds
(without alignments)
2919.508 Million cell updates/sec

Title: US-09-680-514-4
Perfect score: 1047
Sequence: 1 ATGCGAGCTGACTGAATTGCT.....TACGCCACTTCCCGAGCCC 1047

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues
Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	526.6	50.3	1059	3	US-08-471-045-57 Sequence 57, Appl
2	526.6	50.3	1059	3	US-08-469-712A-57 Sequence 57, Appl
3	526.6	50.3	1062	1	US-08-347-029-3 Sequence 3, Appli
4	526.6	50.3	1062	2	US-08-457-254-3 Sequence 3, Appli
5	526.6	50.3	1062	2	US-08-484-257-18 Sequence 18, Appl
6	526.6	50.3	1062	3	US-08-482-212-1 Sequence 1, Appli
7	526.6	50.3	1062	3	US-08-999-927-1 Sequence 1, Appli
8	526.6	50.3	1062	5	PCT-US94-08806-18 Sequence 18, Appl
9	526.6	50.3	1062	5	PCT-US95-01829-1 Sequence 1, Appli
10	526.6	50.3	1062	5	PCT-US95-14932-3 Sequence 3, Appli
11	526.6	50.3	1062	5	PCT-US95-16626-3 Sequence 3, Appli
12	526.6	50.3	1342	1	US-08-388-779A-1 Sequence 1, Appli
13	526.6	50.3	1342	1	US-08-591-070A-1 Sequence 1, Appli
14	526.6	50.3	1342	1	US-08-413-803-24 Sequence 24, Appl
15	526.6	50.3	1342	1	US-08-321-488A-24 Sequence 24, Appl
16	526.6	50.3	1342	2	US-08-414-161B-1 Sequence 1, Appli
17	526.6	50.3	1342	2	US-08-927-855-1 Sequence 1, Appli
18	526.6	50.3	1342	5	PCT-US95-03776-24 Sequence 24, Appl
19	526.6	50.3	1342	5	PCT-US95-03776-26 Sequence 26, Appl
20	526.6	50.3	1795	1	US-08-330-517-1 Sequence 1, Appli
21	525	50.1	605	1	US-08-388-779A-3 Sequence 3, Appli
22	525	50.1	605	1	US-08-591-070A-3 Sequence 3, Appli
23	525	50.1	605	2	US-08-927-855-3 Sequence 3, Appli
24	486.8	46.5	525	1	US-08-434-411-1 Sequence 1, Appli
25	486.8	46.5	525	1	US-08-434-402-1 Sequence 1, Appli
26	486.8	46.5	525	1	US-08-783-288-1 Sequence 1, Appli
27	486.8	46.5	525	2	US-08-890-640-1 Sequence 1, Appli

28	486.8	46.5	525	6	5194592-25 Patent No. 5194592
29	478	45.7	1164	1	US-08-413-803-26 Sequence 26, Appl
30	478	45.7	1164	1	US-08-321-488A-26 Sequence 26, Appl
31	478	45.7	1164	5	PCT-US95-03776-28 Sequence 28, Appl
32	474.6	45.3	546	3	US-08-469-318-177 Sequence 177, App
33	474.6	45.3	546	3	US-08-469-318-177 Sequence 177, App
34	474.6	45.3	546	3	US-08-468-609A-177 Sequence 178, App
35	474.6	45.3	546	3	US-08-468-609A-177 Sequence 178, App
36	474.6	45.3	546	5	PCT-US95-01185-177 Sequence 177, App
37	474.6	45.3	546	5	PCT-US95-01185-177 Sequence 178, App
38	474.6	45.3	921	3	US-08-469-318-72 Sequence 72, Appl
39	474.6	45.3	921	3	US-08-469-318-75 Sequence 75, Appl
40	474.6	45.3	921	3	US-08-469-318-78 Sequence 78, Appl
41	474.6	45.3	921	3	US-08-469-318-84 Sequence 72, Appl
42	474.6	45.3	921	3	US-08-468-609A-72 Sequence 72, Appl
43	474.6	45.3	921	3	US-08-468-609A-75 Sequence 75, Appl
44	474.6	45.3	921	3	US-08-468-609A-78 Sequence 78, Appl
45	474.6	45.3	921	3	US-08-468-609A-84 Sequence 84, Appl

ALIGNMENTS

RESULT 1
US-08-471-045-57
; Sequence 57, Application US/08471045
; Patent No. 6060047
; GENERAL INFORMATION:
; APPLICANT: Abrams, Mark A.
; APPLICANT: Bauer, S. C.
; APPLICANT: Braford-Goldberg, Sarah R.
; APPLICANT: Caparon, Mair H.
; APPLICANT: Easton, Alan M.
; APPLICANT: Klein, Barbara K.
; APPLICANT: McKearn, John P.
; APPLICANT: Olin, Peter O.
; APPLICANT: Paik, Kumnan
; APPLICANT: Thomas, John W.
; TITLE OF INVENTION: Co-administration of Interleukin-3
; TITLE OF INVENTION: Mutant Polypeptides with CSF's for Multi-lineage
; TITLE OF INVENTION: Hematopoietic Cell Production
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,
; ADDRESSEE: Corporate Patent Dept.
; STREET: P. O. Box 5110
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60680
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,045
; FILING DATE: 06-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/01184
; FILING DATE: 02-FEB-1995
; PRIOR APPLICATION DATA: US 08/193,373
; APPLICATION NUMBER: 04-FEB-1994
; FILING DATE: 04-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Dennis A.
; REGISTRATION NUMBER: 34,547
; REFERENCE/DOCKET NUMBER: C-2789/3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (708)470-6501
; TELEFAX: (708)470-6881
; INFORMATION FOR SEQ ID NO: 57:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1059 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-471-045-57

```

```

Query Match          50.3%; Score 526.6; DB 3; Length 1059;
Best Local Similarity 99.2%; Pred. No. 2.1e-119;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATTCCTCGTGGTCACTCTCTCTAACTGCAAGGCTAACGCTG 60
Db 1 ATGGAGCTGACTGAATTCCTCGTGGTCACTCTCTCTAACTGCAAGGCTAACGCTG 60
QY 61 TCCAGCCCGCTCTCTCTGCTGTGACCTCCGAGTCCAGTCCAGTAACTGCTTCTG 120
Db 61 TCCAGCCCGCTCTCTCTGCTGTGACCTCCGAGTCCAGTCCAGTAACTGCTTCTG 120
QY 121 CATGCTCTTACAGCAGACTGAGCCAGTCCGAGAGTTCACCTTTCCTACACCTGTC 180
Db 121 CATGCTCTTACAGCAGACTGAGCCAGTCCGAGAGTTCACCTTTCCTACACCTGTC 180
QY 181 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Db 181 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 241 GCACAGGACATTCCTGGAGCAGTACCTTCTGCTGGAGGAGTGGAGGAGGAGGAG 300
Db 241 GCACAGGACATTCCTGGAGCAGTACCTTCTGCTGGAGGAGTGGAGGAGGAGGAG 300
QY 301 CAACTGGGAGCCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 301 CAACTGGGAGCCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 361 CTCCTTGGGGCCCTCGACAGCTTCCTTGGAACTCCCTCCACAGGCGGAGGAGGAG 420
Db 361 CTCCTTGGGGCCCTCGACAGCTTCCTTGGAACTCCCTCCACAGGCGGAGGAGGAG 420
QY 421 GCTCACAAGGATCCCAATGCCATCTCTCTGAGCTTCCACACCTGCTCCGAGGAAAGTG 480
Db 421 GCTCACAAGGATCCCAATGCCATCTCTCTGAGCTTCCACACCTGCTCCGAGGAAAGTG 480
QY 481 CGTTTCTCTGATGCTGTAGGAGGGTCCACCTCTCGGTAGCGGGGCGGCCAAC 533
Db 481 CGTTTCTCTGATGCTGTAGGAGGGTCCACCTCTCGGTAGCGGGGCGGCCAAC 533

```

RESULT 2

```

US-08-469-712A-57
; Sequence 57, Application US/08469712A
; Patent No. 6093395
; GENERAL INFORMATION:
; APPLICANT: Abrams, Mark A.
; APPLICANT: Bauer, S. C.
; APPLICANT: Brarford-Goldberg, Sarah R.
; APPLICANT: Caparon, Mair H.
; APPLICANT: Easton, Alan M.
; APPLICANT: Klein, Barbara K.
; APPLICANT: McKearn, John P.
; APPLICANT: Ollins, Peter O.
; APPLICANT: Paik, Kumman
; APPLICANT: Thomas, John W.
; TITLE OF INVENTION: Co-administration of Interleukin-3
; TITLE OF INVENTION: Mutant Polypeptides with CSP's for Multi-lineage
; TITLE OF INVENTION: Hematopoietic Cell Production
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,
; ADDRESSEE: Corporate Patent Dept.
; STREET: P. O. Box 5110

```

```

; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60680
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,712A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/011184
; FILING DATE: 02-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/193,373
; FILING DATE: 04-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Dennis A.
; REGISTRATION NUMBER: 34,547
; REFERENCE/DOCKET NUMBER: C-2789/2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (708)470-8501
; TELEFAX: (708)470-6881
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1059 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-469-712A-57

```

```

Query Match          50.3%; Score 526.6; DB 3; Length 1059;
Best Local Similarity 99.2%; Pred. No. 2.1e-119;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
Db 1 ATGGAGCTGACTGAATTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
QY 61 TCCAGCCCGCTCTCTCTGCTGTGACCTCCGAGTCCAGTCCAGTAACTGCTTCTG 120
Db 61 TCCAGCCCGCTCTCTCTGCTGTGACCTCCGAGTCCAGTCCAGTAACTGCTTCTG 120
QY 121 CATGCTCTTACAGCAGACTGAGCCAGTCCGAGAGTTCACCTTTCCTACACCTGTC 180
Db 121 CATGCTCTTACAGCAGACTGAGCCAGTCCGAGAGTTCACCTTTCCTACACCTGTC 180
QY 181 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Db 181 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 241 GCACAGGACATTCCTGGAGCAGTACCTTCTGCTGGAGGAGTGGAGGAGGAGGAG 300
Db 241 GCACAGGACATTCCTGGAGCAGTACCTTCTGCTGGAGGAGTGGAGGAGGAGGAG 300
QY 301 CAACTGGGAGCCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 301 CAACTGGGAGCCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 361 CTCCTTGGGGCCCTCGACAGCTTCCTTGGAACTCCCTCCACAGGCGGAGGAGGAG 420
Db 361 CTCCTTGGGGCCCTCGACAGCTTCCTTGGAACTCCCTCCACAGGCGGAGGAGGAG 420
QY 421 GCTCACAAGGATCCCAATGCCATCTCTCTGAGCTTCCACACCTGCTCCGAGGAAAGTG 480
Db 421 GCTCACAAGGATCCCAATGCCATCTCTCTGAGCTTCCACACCTGCTCCGAGGAAAGTG 480
QY 481 CGTTTCTCTGATGCTGTAGGAGGGTCCACCTCTCGGTAGCGGGGCGGCCAAC 533

```

Db 481 CGTTTCTGATGCTTGTAGGAGGTCACCCCTCTGCTCAGCGGGCCCCACC 533
|||||

RESULT 3
US-08-347-029-3
; Sequence 3, Application US/08347029
; Patent No. 5641655
; GENERAL INFORMATION:
; APPLICANT: Foster, Donald C.
; APPLICANT: Heipel, Mark
; APPLICANT: Holly, Richard D.
; TITLE OF INVENTION: METHOD FOR SECRETING THROMBOPOIETIN
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/347,029
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-648
; REFERENCE/DOCKET NUMBER: 94-13
; TELEPHONE: 206-442-6673
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1062 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1059
US-08-347-029-3

Query Match 50.3%; Score 526.6; DB 1; Length 1062;
Best Local Similarity 99.2%; Pred. No. 2.1e-119;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ATGAGCTGACTGAATTCCTCGTGGTTCATGCTTCTCTAACTCAAGCTAACGCTG 60
|||||
Db 1 ATGAGCTGACTGAATTCCTCGTGGTTCATGCTTCTCTAACTCAAGCTAACGCTG 60
|||||
QY 61 TCAGCCCGGCTCCTCCTGCTTGTGACCTCCGAGTCCCTCAGTAACTGCTTCGTGACTCC 120
|||||
Db 61 TCAGCCCGGCTCCTCCTGCTTGTGACCTCCGAGTCCCTCAGTAACTGCTTCGTGACTCC 120
|||||
QY 121 CATGTCCTTCACAGCACTGACCCAGTCCAGAGGTTACCCCTTTGCTACACCTGTC 180
|||||
Db 121 CATGTCCTTCACAGCACTGACCCAGTCCAGAGGTTACCCCTTTGCTACACCTGTC 180
|||||
QY 181 CTGCTCCCTGCTGTGACTTTAGCTTGGGGAATGAAACCAGATGGAGGACCAAG 240
|||||
Db 181 CTGCTCCCTGCTGTGACTTTAGCTTGGGGAATGAAACCAGATGGAGGACCAAG 240
|||||
QY 241 GCACAGACATTCCTGGAGCAGTGACCCCTTCCTGCTGGAGGAGTGTATGCGACGCGGA 300
|||||

Db 241 GCACAGACATTCCTGGAGCAGTGACCCCTTCCTGCTGGAGGAGTGTATGCGACGACGCGGA 300
QY 301 CAATGGGACCCACTTGCCTCTCATCCCTCTGCGGGCAGCTTTCTGGACAGGTCCTGCTC 360
|||||
Db 301 CAATGGGACCCACTTGCCTCTCATCCCTCTGCGGGCAGCTTTCTGGACAGGTCCTGCTC 360
|||||
QY 361 CTCCTTGGGGCCCTGTCAGAGCCCTTTGGAACCCAGCTTCTCCACAGGCGAGGACCA 420
|||||
Db 361 CTCCTTGGGGCCCTGTCAGAGCCCTTTGGAACCCAGCTTCTCCACAGGCGAGGACCA 420
|||||
QY 421 GCTCACAAGGATCCCAATGCCATCTTCCTGAGCTTCAACACCTGCTCCAGGAAAGGTG 480
|||||
Db 421 GCTCACAAGGATCCCAATGCCATCTTCCTGAGCTTCAACACCTGCTCCAGGAAAGGTG 480
|||||
QY 481 CGTTTCTGATGCTTGTAGGAGGTCACCCCTCTGCTACGGGGGGGCCAAC 533
|||||
Db 481 CGTTTCTGATGCTTGTAGGAGGTCACCCCTCTGCTACGGGGGGGCCAAC 533
|||||

RESULT 4
US-08-457-254-3
; Sequence 3, Application US/08457254
; Patent No. 5986049
; GENERAL INFORMATION:
; APPLICANT: Forstrom, John W
; APPLICANT: Lofton-Day, Catherine E
; APPLICANT: Lok, Si
; TITLE OF INVENTION: PURIFIED THROMBOPOIETIN AND METHOD OF
; TITLE OF INVENTION: MAKING IT
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,254
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31,648
; REFERENCE/DOCKET NUMBER: 94-11C1
; TELEPHONE: 206-442-6673
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1062 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1059
US-08-457-254-3

Query Match 50.3%; Score 526.6; DB 2; Length 1062;
Best Local Similarity 99.2%; Pred. No. 2.1e-119;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ATGAGCTGACTGAATTCCTCGTGGTTCATGCTTCTCTAACTCAAGCTAACGCTG 60
|||||
Db 1 ATGAGCTGACTGAATTCCTCGTGGTTCATGCTTCTCTAACTCAAGCTAACGCTG 60
|||||

```
QY 61 TCAGAGCCGGCCCTCCTGCTGTGACCTCCGAGTCTCAGTAACTGCTTGTGACTCC 120
Db 61 TCAGAGCCGGCCCTCCTGCTGTGACCTCCGAGTCTCAGTAACTGCTTGTGACTCC 120
QY 121 CATGCTCTTCACAGAGACTGAGCCAGTGCACAGGTTTCCCTTACACCTGTC 180
Db 121 CATGCTCTTCACAGAGACTGAGCCAGTGCACAGGTTTCCCTTACACCTGTC 180
QY 181 CTGCTGCCTGCTGTGACTTTAGCTTGGAGATGGAAACCCAGATGGAGGAGACCAAG 240
Db 181 CTGCTGCCTGCTGTGACTTTAGCTTGGAGATGGAAACCCAGATGGAGGAGACCAAG 240
QY 241 GCACAGGACATCTGGGAGCAGTCAACCTTCTGCTGGAGGAGTGTGAGCAGCAGGGGA 300
Db 241 GCACAGGACATCTGGGAGCAGTCAACCTTCTGCTGGAGGAGTGTGAGCAGCAGGGGA 300
QY 301 CAACCTGGGAGCCACTGCTCTCATCCCTGCTGGGGAGCTTCTGACAGGTCCTGTC 360
Db 301 CAACCTGGGAGCCACTGCTCTCATCCCTGCTGGGGAGCTTCTGACAGGTCCTGTC 360
QY 361 CTCTTGGGGCCCTGCAGAGCTCCTTGGAAACCCAGCTTCTCCACAGGCGAGGACCA 420
Db 361 CTCTTGGGGCCCTGCAGAGCTCCTTGGAAACCCAGCTTCTCCACAGGCGAGGACCA 420
QY 421 GCTCACAAGGATCCATGCAATGCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGTG 480
Db 421 GCTCACAAGGATCCATGCAATGCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGTG 480
QY 481 CGTTTCTGATGCTTGTAGGAGGTCACACCTCTGGTAGCGGGGGCGCCCAAC 533
Db 481 CGTTTCTGATGCTTGTAGGAGGTCACACCTCTGGTAGCGGGGGCGCCCAAC 533
```

RESULT 5

```
US-08-484-257-18
; Sequence 18, Application US/08484257
; Patent No. 5985537
```

GENERAL INFORMATION:

```
; APPLICANT: Holly, Richard D.
; APPLICANT: Lok, Si
; APPLICANT: Foster, Donald C.
; APPLICANT: Hagen, Frederick S.
; APPLICANT: Kaushansky, Kenneth
; APPLICANT: Kuiper, Joseph L.
; APPLICANT: Lofton-Day, Catherine E.
; APPLICANT: Oort, Pieter J.
```

```
; TITLE OF INVENTION: Methods for stimulating Granulocyte/Macrophage Lineage Cells
```

```
; NUMBER OF SEQUENCES: 21
```

```
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: ZymoGenetics, Inc.
```

```
; STREET: 1201 Eastlake E.
```

```
; CITY: Seattle
```

```
; STATE: WA
```

```
; COUNTRY: USA
```

```
; ZIP: 98102
```

```
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: Floppy disk
```

```
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/08/484,257
```

```
; FILING DATE:
```

```
; CLASSIFICATION: 514
```

```
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Parker, Gary E
```

```
; REGISTRATION NUMBER: 31-648
```

```
; REFERENCE/DOCKET NUMBER: 95-09
```

```
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: 206-442-6673
```

```
; TELEFAX: 206-442-6678
```

```
; INFORMATION FOR SEQ ID NO: 18:
```

; SEQUENCE CHARACTERISTICS:

```
; LENGTH: 1062 base pairs
```

```
; TYPE: nucleic acid
```

```
; STRANDEDNESS: double
```

```
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: cDNA
```

```
; FEATURE:
```

```
; NAME/KEY: CDS
```

```
; LOCATION: 1..1059
```

```
; US-08-484-257-18
```

Query Match

```
Best Local Similarity 50.3%; Score 526.6; DB 2; Length 1062;
```

```
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 ATGGAGCTGACTGAATTGCTCTCGTGGTCATGCTTCTCTAACTCAAGGCTAAAGCTG 60
Db 1 ATGGAGCTGACTGAATTGCTCTCGTGGTCATGCTTCTCTAACTCAAGGCTAAAGCTG 60
QY 61 TCAGAGCCGGCCCTCCTGCTGTGACCTCCGAGTCTCAGTAACTGCTTGTGACTCC 120
Db 61 TCAGAGCCGGCCCTCCTGCTGTGACCTCCGAGTCTCAGTAACTGCTTGTGACTCC 120
QY 121 CATGCTCTTCACAGAGACTGAGCCAGTGCACAGGTTTCCCTTACACCTGTC 180
Db 121 CATGCTCTTCACAGAGACTGAGCCAGTGCACAGGTTTCCCTTACACCTGTC 180
QY 181 CTGCTGCCTGCTGTGACTTTAGCTTGGAGATGGAAACCCAGATGGAGGAGACCAAG 240
Db 181 CTGCTGCCTGCTGTGACTTTAGCTTGGAGATGGAAACCCAGATGGAGGAGACCAAG 240
QY 241 GCACAGGACATCTGGGAGCAGTCAACCTTCTGCTGGAGGAGTGTGAGCAGCAGGGGA 300
Db 241 GCACAGGACATCTGGGAGCAGTCAACCTTCTGCTGGAGGAGTGTGAGCAGCAGGGGA 300
QY 301 CAACCTGGGAGCCACTGCTCTCATCCCTGCTGGGGAGCTTCTGACAGGTCCTGTC 360
Db 301 CAACCTGGGAGCCACTGCTCTCATCCCTGCTGGGGAGCTTCTGACAGGTCCTGTC 360
QY 361 CTCTTGGGGCCCTGCAGAGCTCCTTGGAAACCCAGCTTCTCCACAGGCGAGGACCA 420
Db 361 CTCTTGGGGCCCTGCAGAGCTCCTTGGAAACCCAGCTTCTCCACAGGCGAGGACCA 420
QY 421 GCTCACAAGGATCCATGCAATGCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGTG 480
Db 421 GCTCACAAGGATCCATGCAATGCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGTG 480
QY 481 CGTTTCTGATGCTTGTAGGAGGTCACACCTCTGGTAGCGGGGGCGCCCAAC 533
Db 481 CGTTTCTGATGCTTGTAGGAGGTCACACCTCTGGTAGCGGGGGCGCCCAAC 533
```

RESULT 6

```
US-08-482-212-1
```

```
; Sequence 1, Application US/08482212
```

```
; Patent No. 6013067
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Fibbe, Willem E.
```

```
; APPLICANT: Grossmann, Angelika
```

```
; TITLE OF INVENTION: METHODS FOR INCREASING HEMATOPOIETIC CELLS
```

```
; NUMBER OF SEQUENCES: 2
```

```
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: ZymoGenetics, Inc.
```

```
; STREET: 1201 Eastlake Avenue East
```

```
; CITY: Seattle
```

```
; STATE: WA
```

```
; COUNTRY: USA
```

```
; ZIP: 98102
```

```
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: Floppy disk
```

```
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
```



```

; TITLE OF INVENTION: Using Hematopoietic Proteins.
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/999,927
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/347,748
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-648
; REFERENCE/DOCKET NUMBER: 94-09C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6673
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1062 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1059
; US-08-999-927-1

Query Match 50.3%; Score 526.6; DB 3; Length 1062;
Best Local Similarity 99.2%; Pred. No. 2.le-119;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps

QY 1 ATGGAGCTGACTGAATTGCTCTCGTGGTCACTGTTCTCTCTAACTCAAGGCTAACGCTG 60
Db 1 ATGGAGCTGACTGAATTGCTCTCGTGGTCACTGTTCTCTCTAACTCAAGGCTAACGCTG 60
QY 61 TCCAGCCCGGCTCCTCTGCTGTTGACCTCCGAGTCTCTAGTAACTGCTTCGTGACTCC 120
Db 61 TCCAGCCCGGCTCCTCTGCTGTTGACCTCCGAGTCTCTAGTAACTGCTTCGTGACTCC 120
QY 121 CATGTCTTCACAGCAGACTGAGCCAGTGCACAGAGTTGACCCCTTTGCTACACCTGTC 180
Db 121 CATGTCTTCACAGCAGACTGAGCCAGTGCACAGAGTTTCAACCTTTGCTACACCTGTC 180
QY 181 CTGCTGCCTGCTGTGACTTTAGCTTGGGAGATGGAACACCCAGATGGAGGAGACCAAG 240
Db 181 CTGCTGCCTGCTGTGACTTTAGCTTGGGAGATGGAACACCCAGATGGAGGAGACCAAG 240
QY 241 GCACAGCACATTCTGGGAGCAGTGACCTTCTGCTGGAGGAGTGATGGCAGACGCGGA 300
Db 241 GCACAGCACATTCTGGGAGCAGTGACCTTCTGCTGGAGGAGTGATGGCAGACGCGGA 300
QY 301 CAACATGGGACCCACTTGCCCTCTCATCCCTCTGGGAGAGCTTTCTTGGACAGGTCCTCTC 360
Db 301 CAACATGGGACCCACTTGCCCTCTCATCCCTCTGGGAGAGCTTTCTTGGACAGGTCCTCTC 360
QY 361 CTCTTTGGGGCCCTGCAGAGCCTCTTGGAAACCAGCTTCTCCACAGGGCAGGACCACA 420
Db 361 CTCTTTGGGGCCCTGCAGAGCCTCTTGGAAACCAGCTTCTCCACAGGGCAGGACCACA 420

```



```
; LOCATION: 1..1059
PCT-US95-01829-1

Query Match          50.3%; Score 526.6; DB 5; Length 1062;
Best Local Similarity 99.2%; Pred. No. 2.1e-119;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAGCTGACTGAATTCCTCGTGGTGCATGCTTCTTAAGTCAAGGCTAACGCTG 60
   |||||||
Db 1 ATGAGCTGACTGAATTCCTCGTGGTGCATGCTTCTTAAGTCAAGGCTAACGCTG 60

QY 61 TCAGAGCCGGCTCCTCGTGTGACCTCCGAGTCCAGTAACTGCTTCGTGACTCC 120
   |||||||
Db 61 TCAGAGCCGGCTCCTCGTGTGACCTCCGAGTCCAGTAACTGCTTCGTGACTCC 120

QY 121 CATGTCTTCACAGCAGACTGAGCCAGTCCAGAGGTTTACCCCTTTCACCTGTC 180
   |||||||
Db 121 CATGTCTTCACAGCAGACTGAGCCAGTCCAGAGGTTTACCCCTTTCACCTGTC 180

QY 181 CTGCTGCTGCTGCTGGAGACTTTCAGTGGAGATGGAACCCAGATGGAGGACCAAG 240
   |||||||
Db 181 CTGCTGCTGCTGCTGGAGACTTTCAGTGGAGATGGAACCCAGATGGAGGACCAAG 240

QY 241 GCACAGGACATTCCTGGAGCAGTGACCCCTTCTGCTGGAGGAGTGATGGCAGCAGGGGA 300
   |||||||
Db 241 GCACAGGACATTCCTGGAGCAGTGACCCCTTCTGCTGGAGGAGTGATGGCAGCAGGGGA 300

QY 301 CAACTGGGACCCACTTGCCTCTCATCCCTCTGGGGAGCTTCTTGACAGAGTCCGCTTC 360
   |||||||
Db 301 CAACTGGGACCCACTTGCCTCTCATCCCTCTGGGGAGCTTCTTGACAGAGTCCGCTTC 360

QY 361 CTCTTTGGGCGCTTGCAGAGCCTTCTTGGAAACCCAGCTTCTCCACAGGCGAGGACCA 420
   |||||||
Db 361 CTCTTTGGGCGCTTGCAGAGCCTTCTTGGAAACCCAGCTTCTCCACAGGCGAGGACCA 420

QY 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTTCTCCGAGGAAAGGTG 480
   |||||||
Db 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTTCTCCGAGGAAAGGTG 480

QY 481 CGTTTCTGTATGCTTGTAGGAGGTCACCCCTCTGCTAGCGGGGGGGCCCAAC 533
   |||||||
Db 481 CGTTTCTGTATGCTTGTAGGAGGTCACCCCTCTGCTAGCGGGGGGGCCCAAC 533

RESULT 11
PCT-US95-16626-3
; Sequence 3, Application PC/TUS9516626
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; APPLICANT: 1201 Eastlake Avenue East
; APPLICANT: Seattle
; APPLICANT: WA
; APPLICANT: USA
; APPLICANT: 98102
; TITLE OF INVENTION: PURIFIED THROMBOPOIETIN AND METHOD OF
```

```

: TITLE OF INVENTION: MAKING IT
: NUMBER OF SEQUENCES: 24
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: zymoGenetics, Inc.
: STREET: 1201 Eastlake Avenue East
: CITY: Seattle
: STATE: WA
: COUNTRY: USA
: ZIP: 98102
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/16626
: FILING DATE:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Parker, Gary E
: REGISTRATION NUMBER: 31,648
: REFERENCE/DOCKET NUMBER: 94-11PC
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 206-442-6673
: TELEFAX: 206-442-6678
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1062 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: CDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1059
: PCT-US95-16626-3

```

Query Match	50.3%	Score 526.6;	DB 5;	Length 1062;
Best Local Similarity	99.2%;	Mism. No. 2.le-119;		
Matches 529;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;
QY	1	ATGGAGCTGACTGAATTCCTCCTCGTGGTGCATGCTTCTCCTAACTCAAGAGCTAAAGCTG	60	
Db	1	ATGGAGCTGACTGAATTCCTCCTCGTGGTGCATGCTTCTCCTAACTCAAGAGCTAAAGCTG	60	
QY	61	TCACGCCGGCTCCTCCGCTGTTGTGACCTCCGAGTCCCTCAGTAAACTGCTTCGTGACTCC	120	
Db	61	TCACGCCGGCTCCTCCTGCTTGTGTGACCTCCGAGTCCCTCAGTAAACTGCTTCGTGACTCC	120	
QY	121	CATGTCTCTTACAGCAGACTGAGCCAGTGCACAGAGGTTACCCCTTTCGCTACACCTGTC	180	
Db	121	CATGTCTCTTACAGCAGACTGAGCCAGTGCACAGAGGTTACCCCTTTCGCTACACCTGTC	180	
QY	181	CTGCTGCCTGCTGTGGACTTTAGCTTGGGAGANTGAAACCCAGATGGAGGAGACCAG	240	
Db	181	CTGCTGCCTGCTGTGGACTTTAGCTTGGGAGANTGAAACCCAGATGGAGGAGACCAG	240	
QY	241	GCACAGGACATCTCTGGAGCAGTGCACCTTCCTGCTGGAGGAGTGATGCCACACGGGA	300	
Db	241	GCACAGGACATCTCTGGAGCAGTGCACCTTCCTGCTGGAGGAGTGATGCCACACGGGA	300	
QY	301	CAACTGGAGCCACTTGCTCTCATCCCTCCTGGGGAGCTTTCTGGACAGGTCGGTCTC	360	
Db	301	CAACTGGAGCCACTTGCTCTCATCCCTCCTGGGGAGCTTTCTGGACAGGTCGGTCTC	360	
QY	361	CTCCTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGTTCCTCCACAGGGCAGGACCACA	420	
Db	361	CTCCTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGTTCCTCCACAGGGCAGGACCACA	420	
QY	421	GCTCAAGAGTCCCAATGCCATCTTCCTGAGCTTCCCAACACTGCTCCGAGGAAAGTG	480	
Db	421	GCTCAAGAGTCCCAATGCCATCTTCCTGAGCTTCCCAACACTGCTCCGAGGAAAGTG	480	

```

Qy 481 CGTTTCCTGATGCTTGTAGGAGGTCCACCCCTCTCGGTACGCGGGGCGCCAAC 533
      |||||
Db 481 CGTTTCCTGATGCTTGTAGGAGGTCCACCCCTCTCGGTACGCGGGGCGCCAAC 533
      |||||

RESULT 12
US-08-388-779A-1
; Sequence 1, Application US/08388779A
; Patent No. 5696250
; GENERAL INFORMATION:
; APPLICANT: Elliott, Steven G.
; TITLE OF INVENTION: MGDF ANALOGS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1720
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/388,779A
; APPLICATION NUMBER: US/08/388,779A
; FILING DATE: 15-FEB-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Cook Ph.D., Robert R.
; REGISTRATION NUMBER: 31,602
; REFERENCE/DOCKET NUMBER: A-337
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1342 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36..1094
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 99..1094
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 36..98
US-08-388-779A-1

```

Query Match	50.3%;	Score	526.6;	DB 1:	Length	1342;
Best Local Similarity	99.2%;	Pred. No.	2.2e-119;			
Matches 529;	Conservative	0;	Mismatches	4;	Indels	0;
Gaps						
QY	1	ATGGAGCTGACTGAATTGCTCCTCGTGGTCACTGCTTCTTAACTCGAAGGCTAAGCGTG	60			
DB	36	ATGGAGCTGACTGAATTGCTCCTCGTGGTCACTGCTTCTTAACTCGAAGGCTAAGCGTG	95			
QY	61	TCAGAGCCGGCTCCTCCTGTTGTGACCTCCGAGTCCCTCAGTAAACTGCTTCGTGACTCC	120			
DB	96	TCAGAGCCGGCTCCTCCTGTTGTGACCTCCGAGTCCCTCAGTAAACTGCTTCGTGACTCC	155			
QY	121	CATGTCCTTTCACAGCAGACTGAGCCAGTCCCCAGAGGTTACCCCTTGCTTACACCTGTC	180			
DB	156	CATGTCCTTTCACAGCAGACTGAGCCAGTCCCCAGAGGTTACCCCTTGCTTACACCTGTC	215			
QY	181	CTGCTGCCCTGCTGTGGACTTTTAGCTTGGGAAATGGAAACCAGATGGAGGAGACCAAG	240			
DB	216	CTGCTGCCCTGCTGTGGACTTTTAGCTTGGGAAATGGAAACCAGATGGAGGAGACCAAG	275			

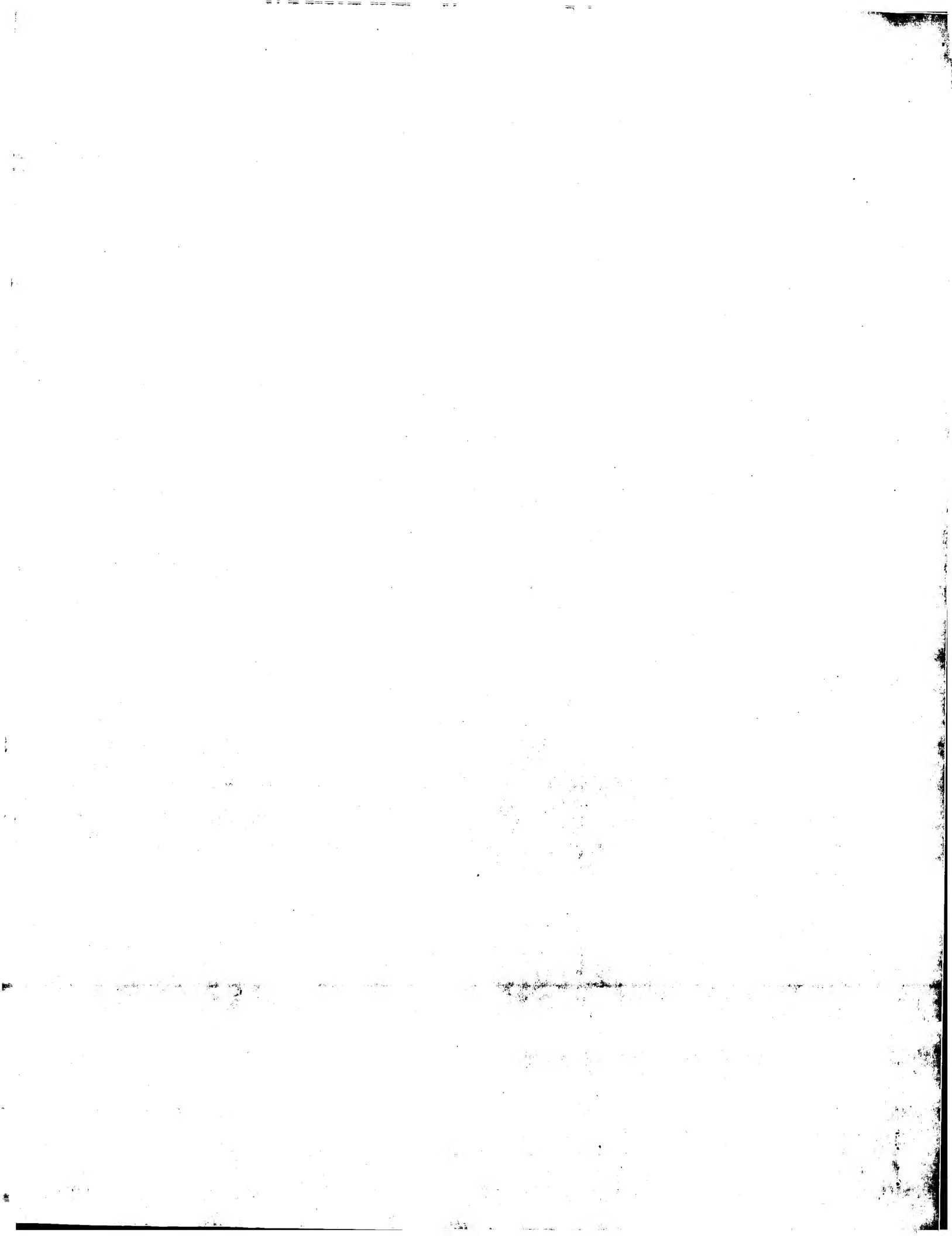
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/252,628
;; FILING DATE: 31-MAY-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/321,488
;; FILING DATE: 12-OCT-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/347,780
;; FILING DATE: 30-NOV-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Cook Ph.D., Robert R.
;; REGISTRATION NUMBER: 31,602
;; REFERENCE/DOCKET NUMBER: A-290B
;; INFORMATION FOR SEQ ID NO: 24:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1342 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: unknown
;; TOPOLOGY: unknown
;; MOLECULE TYPE: CDNA
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 36..1097
;; FEATURE:
;; NAME/KEY: mat_peptide
;; LOCATION: 99..1097
;; FEATURE:
;; NAME/KEY: sig_peptide
;; LOCATION: 36..98
US-08-413-803-24

Query Match 50.3%; Score 526.6; DB 1; Length 1342;
Best Local Similarity 99.2%; Pred. No. 2.2e-119;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ATGAGCTGACTGAATTCCTCGTGGTCTATGCTTCTCTACTCAAGGCTAAGCTG 60
Db |||||||
QY 36 ATGAGCTGACTGAATTCCTCGTGGTCTATGCTTCTCTACTCAAGGCTAAGCTG 95
Db |||||||
QY 61 TCCAGCCGGGCTCTCTGCTGTGTGACCTCCGAGTCCCTCAGTAACTGCTTGTGACTCC 120
Db |||||||
QY 96 TCCAGCCGGGCTCTCTGCTGTGTGACCTCCGAGTCCCTCAGTAACTGCTTGTGACTCC 155
Db |||||||
QY 121 CATGCTCTTCCACAGCAGCTGAGCCAGTCCCGAGGTTTCACTTGTGCTACCTGTC 180
Db |||||||
QY 156 CATGCTCTTCCACAGCAGCTGAGCCAGTCCCGAGGTTTCACTTGTGCTACCTGTC 215
Db |||||||
QY 181 CTGCTGCTGCTGTGGACTTTAGCTTGGGAGATGAAACCCAGATGGAGGACCAAG 240
Db |||||||
QY 216 CTGCTGCTGCTGTGGACTTTAGCTTGGGAGATGAAACCCAGATGGAGGACCAAG 275
Db |||||||
QY 241 GCACAGACATCTGGGAGCAGTACCTTCTGCTGGAGGAGTATGACGACGAGGGA 300
Db |||||||
QY 276 GCACAGACATCTGGGAGCAGTACCTTCTGCTGGAGGAGTATGACGACGAGGGA 335
Db |||||||
QY 301 CAACCTGGGACCCACTTGGCTCTCATCCCTCTGGGAGGCTTCTGGAGAGTCTGCTC 360
Db |||||||
QY 336 CAACCTGGGACCCACTTGGCTCTCATCCCTCTGGGAGGCTTCTGGAGAGTCTGCTC 395
Db |||||||
QY 361 CTCTTTGGGCGCTTCAGAGCCTCTTGGAAACCCAGCTTCTTCCACAGGCGGACCA 420
Db |||||||
QY 396 CTCTTTGGGCGCTTCAGAGCCTCTTGGAAACCCAGCTTCTTCCACAGGCGGACCA 455
Db |||||||
QY 421 GCTCAAGGATCCCAATGCCATCTTCTGAGCTTCCAACTCTGCTCCGAGGAAAGTG 480
Db |||||||
QY 456 GCTCAAGGATCCCAATGCCATCTTCTGAGCTTCCAACTCTGCTCCGAGGAAAGTG 515
Db |||||||
QY 481 CGTTTCTGATGCTGTAGAGGGTCCACCTCTGCTGCTACGCGGCGGCCAAC 533
Db |||||||
QY 516 CGTTTCTGATGCTGTAGAGGGTCCACCTCTGCTGCTACGCGGCGGCCAAC 568
Db |||||||

RESULT 15
US-08-321-488A-24
;; Sequence 24, Application US/08321488A
;; Patent No. 5795569
;; GENERAL INFORMATION:
;; APPLICANT: Bartley, Timothy D.
;; APPLICANT: Bogenberger, Jakob M.
;; APPLICANT: Bosseiman, Robert A.
;; APPLICANT: Hunt, Pamela
;; APPLICANT: Kinstler, Olaf B.
;; APPLICANT: Samal, Babru B.
;; TITLE OF INVENTION: MONO-PGTYLATED PROTEINS THAT STIMULATE
;; TITLE OF INVENTION: MEGAKARYOCYTE GROWTH AND DIFFERENTIATION
;; NUMBER OF SEQUENCES: 27
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: AMGEN INC.
;; STREET: 1840 DeHavilland Drive
;; CITY: Thousand Oaks
;; STATE: California
;; COUNTRY: US
;; ZIP: 91320-1789
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/321,488A
;; FILING DATE: 12-OCT-1994
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/252,628
;; FILING DATE: 31-MAY-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/221,768
;; FILING DATE: 31-MAR-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Cook, Robert R.
;; REGISTRATION NUMBER: 31,602
;; REFERENCE/DOCKET NUMBER: A-290B
;; INFORMATION FOR SEQ ID NO: 24:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1342 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: unknown
;; TOPOLOGY: unknown
;; MOLECULE TYPE: CDNA
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 36..1097
;; FEATURE:
;; NAME/KEY: mat_peptide
;; LOCATION: 99..1097
;; FEATURE:
;; NAME/KEY: sig_peptide
;; LOCATION: 36..98
US-08-321-488A-24

Query Match 50.3%; Score 526.6; DB 1; Length 1342;
Best Local Similarity 99.2%; Pred. No. 2.2e-119;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ATGAGCTGACTGAATTCCTCGTGGTCTATGCTTCTCTACTCAAGGCTAAGCTG 60
Db |||||||
QY 36 ATGAGCTGACTGAATTCCTCGTGGTCTATGCTTCTCTACTCAAGGCTAAGCTG 95
Db |||||||
QY 61 TCCAGCCGGGCTCTCTGCTGTGTGACCTCCGAGTCCCTCAGTAACTGCTTGTGACTCC 120
Db |||||||
QY 96 TCCAGCCGGGCTCTCTGCTGTGTGACCTCCGAGTCCCTCAGTAACTGCTTGTGACTCC 155
Db |||||||
QY 121 CATGCTCTTCCACAGCAGCTGAGCCAGTCCCGAGGTTTCACTTGTGCTACCTGTC 180
Db |||||||

Search completed: April 5, 2002, 10:34:53
Job time: 4987 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2002, 09:10:07 ; Search time 2520.15 Seconds
(without alignments)
6853.777 Million cell updates/sec

Title: US-09-680-514-4
Perfect score: 1047
Sequence: 1 ATGAGCTGACTGAATTGCT.....TACGCCACTTGCCAGCCC 1047

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues
Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.om.*
20: em.or.*
21: em.ov.*
22: em.pat.*
23: em.ph.*
24: em.pl.*
25: em.ro.*
26: em.sts.*
27: em.sy.*
28: em.un.*
29: em.vi.*
30: em.htgo.hum.*
31: em.htgo.inv.*
32: em.htgo.rod.*
33: em.htg.hum.*
34: em.htg.inv.*
35: em.htg.rod.*
36: em.htg.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	526.6	50.3	555	22	E11965	E11965 Human cDNA
2	526.6	50.3	1059	6	AR103959	AR103959 Sequence
3	526.6	50.3	1062	6	AR087133	AR087133 Sequence
4	526.6	50.3	1062	6	AR087571	AR087571 Sequence
5	526.6	50.3	1062	6	I49760	I49760 Sequence 3
6	526.6	50.3	1062	9	HUNTHROMB	L36052 Human throm
7	526.6	50.3	1086	6	E16669	E16669 cDNA encodi
8	526.6	50.3	1267	6	E12182	E12182 Human cDNA
9	526.6	50.3	1342	6	E12214	E12214 Human cDNA
10	526.6	50.3	1342	6	AR008878	AR008878 Sequence
11	526.6	50.3	1342	6	AR023468	AR023468 Sequence
12	526.6	50.3	1342	6	AR087573	AR087573 Sequence
13	526.6	50.3	1342	6	I85129	I85129 Sequence 1
14	526.6	50.3	1691	9	D32047	D32047 Homo sapien
15	526.6	50.3	1721	6	E16668	E16668 cDNA encodi
16	526.6	50.3	1721	22	E11961	E11961 Human cDNA
17	526.6	50.3	1795	6	I33525	I33525 Sequence 1
18	526.6	50.3	1795	9	HUMMLCmpl	L33410 Human c-mpl
19	525	50.1	605	6	AR008879	AR008879 Sequence
20	525	50.1	605	6	AR087574	AR087574 Sequence
21	525	50.1	605	6	I85130	I85130 Sequence 3
22	523.4	50.0	1341	9	HSU11025	U11025 Human megak
23	521.8	49.8	1062	9	HSU59493	U59493 Human throm
24	521.8	49.8	1062	9	HSU59494	U59494 Human throm
25	520.4	49.7	525	6	E02574	E02574 DNA encodin
26	514	49.1	525	6	E02575	E02575 DNA encodin
27	512.4	48.9	525	6	E02576	E02576 DNA encodin
28	492.6	47.0	1050	22	E11559	E11559 Human cDNA
29	486.8	46.5	525	6	AR024358	AR024358 Sequence
30	486.8	46.5	525	6	AR091731	AR091731 Sequence
31	486.8	46.5	525	6	E01731	E01731 cDNA encodi
32	486.8	46.5	525	6	E02573	E02573 DNA encodin
33	486.8	46.5	525	6	E07164	E07164 cDNA encodi
34	486.8	46.5	525	6	E15131	E15131 Human mRNA
35	486.8	46.5	525	6	I71150	I71150 Sequence 1
36	486.8	46.5	525	6	I83709	I83709 Sequence 1
37	486.8	46.5	525	22	E09431	E09431 cDNA encodi
38	486.8	46.5	660	9	HUMGCSF	M17706 Human granu
39	486.8	46.5	1498	9	HSGCSFR1	X03655 Human mRNA
40	486.8	46.5	1521	6	E01219	E01219 cDNA encodi
41	486.8	46.5	1525	6	I05333	I05333 Sequence 11
42	486.8	46.5	1525	6	I08669	I08669 Sequence 1
43	486.8	46.5	1525	6	I09205	I09205 Sequence 5
44	485.2	46.3	522	6	E01631	E01631 DNA encodin
45	485.2	46.3	522	6	E08529	E08529 DNA encodin

ALIGNMENTS

RESULT 1
E11965
ID E11965 standard; RNA; HUM; 555 BP.
XX
AC E11965;
XX
SV E11965.1
XX
DT 08-OCT-1997 (Rel. 52, Created)
DT 02-SEP-2000 (Rel. 65, Last updated, Version 2)
XX
DE Human cDNA encoding thrombopoietin.
XX
KW JP 1996228781-A/6.
XX
OS Homo sapiens (human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX
RN [1]

Db 481 CGTTTCTGATGTTCTAGAGGGTCCACCTCTGCGTCAGCGGGCCCCACC 533

RESULT 3
LOCUS AR087133 1062 bp DNA PAT 07-SEP-2000
DEFINITION Sequence 3 from patent US 5986049.
ACCESSION AR087133
VERSION AR087133.1 GI:10013899
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1062)
AUTHORS Forstrom,J.W., Lofton-Day,C.E. and Lok,S.
TITLE Purified thrombopoietin and method of making it
JOURNAL Patent: US 5986049-A 3 16-NOV-1999;
FEATURES Location/Qualifiers
source 1..1062
BASE COUNT 223 a 375 c 236 g 228 t
ORIGIN

Query Match 50.3%; Score 526.6; DB 6; Length 1062;
Best Local Similarity 99.2%; Pred. No. 1.3e-99;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATGCTCTCGTGGTCATGCTTCTCCTAACTGAGGCTTAAGCGTG 60
|||||
Db 1 ATGGAGCTGACTGAATGCTCTCGTGGTCATGCTTCTCCTAACTGCAAGGCTTAAGCGTG 60
|||||

QY 61 TCCAGCCCGGCTCCTCGTCTGTGACCTCCGAGTCCAGTAACTGCTTCGTGACTCC 120
|||||
Db 61 TCCAGCCCGGCTCCTCGTCTGTGACCTCCGAGTCCAGTAACTGCTTCGTGACTCC 120
|||||

QY 121 CATGCTCTTACACAGCAGACTGAGCCAGTCCAGAGGTTACACCTTTGGCTACACCTGTC 180
|||||
Db 121 CATGCTCTTACACAGCAGACTGAGCCAGTCCAGAGGTTACACCTTTGGCTACACCTGTC 180
|||||

QY 181 CTGCTGCCCTGCTGGAGCAGTGACCTTCTCTGGAGGAGTGATGGCAGCAGCGGGA 240
|||||
Db 181 CTGCTGCCCTGCTGGAGCAGTGACCTTCTCTGGAGGAGTGATGGCAGCAGCGGGA 240
|||||

QY 241 GCACAGGACATCTGGGAGCAGTGACCTTCTCTGGAGGAGTGATGGCAGCAGCGGGA 300
|||||
Db 241 GCACAGGACATCTGGGAGCAGTGACCTTCTCTGGAGGAGTGATGGCAGCAGCGGGA 300
|||||

QY 301 CAACTGGGACCCACTTGGCCTCTCATCCCTCTGGGAGGAGTGATGGCAGCAGCGGGA 360
|||||
Db 301 CAACTGGGACCCACTTGGCCTCTCATCCCTCTGGGAGGAGTGATGGCAGCAGCGGGA 360
|||||

QY 361 CTCTTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACACA 420
|||||
Db 361 CTCTTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACACA 420
|||||

QY 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGGTG 480
|||||
Db 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGGTG 480
|||||

QY 481 CGTTTCTGATGTTCTAGGAGGGTCCACCTCTGCGTACGGGGCGGCCAAC 533
|||||
Db 481 CGTTTCTGATGTTCTAGGAGGGTCCACCTCTGCGTACGGGGCGGCCAAC 533
|||||

RESULT 4
LOCUS AR087571 1062 bp DNA PAT 07-SEP-2000
DEFINITION Sequence 18 from patent US 5989537.
ACCESSION AR087571
VERSION AR087571.1 GI:10014334
KEYWORDS
SOURCE Unknown.

ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1062)
AUTHORS Holly,R.D., Lok,S., Foster,D.C., Hagen,F.S., Kaushansky,K.,
Kuilper,J.L., Lofton-Day,C.E. and Oort,P.J.
TITLE Methods for stimulating granulocyte/macrophage lineage using
thrombopoietin
JOURNAL Patent: US 5989537-A 18 23-NOV-1999;
FEATURES Location/Qualifiers
source 1..1062
BASE COUNT 223 a 375 c 236 g 228 t
ORIGIN

Query Match 50.3%; Score 526.6; DB 6; Length 1062;
Best Local Similarity 99.2%; Pred. No. 1.3e-99;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATGCTCTCGTGGTCATGCTTCTCCTAACTGCAAGGCTTAAGCGTG 60
|||||
Db 1 ATGGAGCTGACTGAATGCTCTCGTGGTCATGCTTCTCCTAACTGCAAGGCTTAAGCGTG 60
|||||

QY 61 TCCAGCCCGGCTCCTCGTCTGTGACCTCCGAGTCCAGTAACTGCTTCGTGACTCC 120
|||||
Db 61 TCCAGCCCGGCTCCTCGTCTGTGACCTCCGAGTCCAGTAACTGCTTCGTGACTCC 120
|||||

QY 121 CATGCTCTTACACAGCAGACTGAGCCAGTCCAGAGGTTACACCTTTGGCTACACCTGTC 180
|||||
Db 121 CATGCTCTTACACAGCAGACTGAGCCAGTCCAGAGGTTACACCTTTGGCTACACCTGTC 180
|||||

QY 181 CTGCTGCCCTGCTGGAGCAGTGACCTTCTCTGGAGGAGTGATGGCAGCAGCGGGA 240
|||||
Db 181 CTGCTGCCCTGCTGGAGCAGTGACCTTCTCTGGAGGAGTGATGGCAGCAGCGGGA 240
|||||

QY 241 GCACAGGACATCTGGGAGCAGTGACCTTCTCTGGAGGAGTGATGGCAGCAGCGGGA 300
|||||
Db 241 GCACAGGACATCTGGGAGCAGTGACCTTCTCTGGAGGAGTGATGGCAGCAGCGGGA 300
|||||

QY 301 CAACTGGGACCCACTTGGCCTCTCATCCCTCTGGGAGGAGTGATGGCAGCAGCGGGA 360
|||||
Db 301 CAACTGGGACCCACTTGGCCTCTCATCCCTCTGGGAGGAGTGATGGCAGCAGCGGGA 360
|||||

QY 361 CTCTTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACACA 420
|||||
Db 361 CTCTTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACACA 420
|||||

QY 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGGTG 480
|||||
Db 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGGTG 480
|||||

QY 481 CGTTTCTGATGTTCTAGGAGGGTCCACCTCTGCGTACGGGGCGGCCAAC 533
|||||
Db 481 CGTTTCTGATGTTCTAGGAGGGTCCACCTCTGCGTACGGGGCGGCCAAC 533
|||||

RESULT 5
LOCUS I49760 1062 bp DNA PAT 07-OCT-1997
DEFINITION Sequence 3 from patent US 5641655.
ACCESSION I49760
VERSION I49760.1 GI:2471980
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1062)
AUTHORS Foster,D.C., Heibel,M.D. and Holly,R.D.
TITLE Methods for producing thrombopoietin polypeptides using a mammalian
tissue plasminogen activator secretory peptide
JOURNAL Patent: US 5641655-A 3 24-JUN-1997;
FEATURES Location/Qualifiers
source 1..1062

BASE COUNT	223 a	375 c	236 g	228 t
ORIGIN	/organism="unknown"			

Query Match 50.3%; Score 526.6; DB 6; Length 1063;
Best Local Similarity 99.2%; Pred. NO. 1.3e-99;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Query Match	50.3%;	Score 526.6;	DB 6;	Length 1062;
Best Local Similarity	99.2%;	Prod No. 12-80		

	Indels	Gaps	0;
Ov	4;	0;	0;

[illegible]

RESULT	6
HUMTHROMB	
LOCUS	1062 bp mRNA
DEFINITION	Human thrombopoietin mRNA, complete cds.
ACCESSION	L36052
VERSION	L36052.1 GI:533216
KEYWORDS	thrombopoietin.
SOURCE	Homo sapiens cDNA to mRNA.
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	
AUTHORS	Foster, D.C., Sprecher, C.A., Grant, F.J., Kramer, J.M., Kuilper, J.L., Holly, R.D., Hartmore, T.E., Heipel, M.D., Bell, L.A.N., Ching, A.F., McGranville, V., Whitcomb, P.J. and Lok, S.
	Human thrombopoietin: gene structure, cDNA sequence, expression, and chromosomal localization
ITILE	
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 91 (26), 13023-13027 (1994)
MEDLINE	95108091

```

FEATURES
  source
    Location/Qualifiers
      1..1062
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
      1..1062
        /note="putative"
        /codon_start=1
        /product="thrombospondin"

```

```
/protein_id="AAC37566.1"
/db_xref="GI:533217"
```

0524861...GI:353217"
/translation="MELTEILLVVMLLTARLTSSPAPACDLVLSKLLRDSHVHL
SRUSQCEVHPHLPVLLPVLAVDFSLGEKWTOMEETKAQDILCAVTLLEGVMAARGOL
GPTCSLLIGLSQGVRLILGQVLLGSLTQLPQPGTAKHDNFNIFTSOHLRLRGKV
RFLMLVGGSTGLVRRAPTTAVPSRTSLVLTNELPNTSGLLETFASRTTSGSL
LKWQCGFRAKIFGLLNQTSRSLDIPGYLNRIHELLNGTRGLPGPSRRTLCAPDIS
GTSDTGSLPNLQPSYRSRSDPPTDQPGVNLHLLNGTLPPLPPTLPVVLHPLLPDPSAPTPP
TSPILNTSYTHSQNLSSQSG"

223 a 375 c 236 g 228 t

BASE COUNT
ORIGIN

	BASE COUNT	ORIGIN
TSPLLNTSYTHSQNLSQEG"	223 a	375 c
	236 g	228 t

Query Match	50.3%	Score 526.6;	DB 9;	Length 1062;
Best Local Similarity	99.2%	Pred. No. 1.3e-99;		
Matches 529;	Conservative	0;	Mismatches 4;	Indels 0; Gaps 0;

[illegible]

RESULT

LOCUS	DEFINITION	REFERENCE	JOURNAL
ACCESSION	AUTHORS	TITLE	COMMENT
VERSION			
KEYWORDS			
SOURCE	ORGANISM		

REFERENCE
1 (bases 1 to 1086)

AUTHORS	Kuroda, K.
TITLE	COMPOSITION CONTAINING THROMBOPOIETIN AND USED FOR STIMULATING RECOVERY OF HEMOSTATIC THROMBUS-FORMING ABILITY OF PATIENT: JP 1998212243-A 3 11-AUG-1998;
JOURNAL	KIRIN BREWERY CO LTD
COMMENT	OS Homo sapiens (human) PN JP 1998212243-A/3

DEFINITION	Human cDNA encoding thrombopoietin.
ACCESSION	E12182
VERSION	E12182.1 GI:3251016
KEYWORDS	JP 1996277296-A/2.
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 1267) Miyazaki, H., Kato, H., Ogami, K., Iwamatsu, A., Akahori, H., Kuroki, R., Shimizu, T. and Muto, T.
TITLE	PROTEIN HAVING TPO ACTIVITY
JOURNAL	Patent: JP 1996277296-A 2 22-OCT-1996;
COMMENT	KIRIN BREWERY CO LTD OS Homo sapiens (human) PN JP 1996277296-A/2 PD 22-OCT-1996 PF 14-FEB-1995 JP 1995161353 PR 14-FEB-1994 JP 94P 39090, 25-MAR-1994 JP 94P 79842, PR 01-JUN-1994 JP 94P 155126, 15-JUN-1994 JP 94P 167328, PR 17-AUG-1994 JP 94P 227159, 17-AUG-1994 JP 94P 193169, PR 01-NOV-1994 JP 94P 304167, 01-DEC-1994 JP 94P 298669, PR 28-DEC-1994 JP 94P 341200 PI MIYAZAKI HIROSHI, KATO HISASHI, OGAMI KINYA, IWAMATSU AKIHIKO, PI AKAHORI HIROMORI, KUROKI RYOTA, SHIMIZU TOSHIYUKI, PI MUTO TAKANORI PC C07K14/52, C07H21/04, C12N15/09, C12P21/02, //A61K38/00, (C12P21/02,

[illegible]

RESULT
E12182

```

Db 325 CAACGGGAGCCACTTCCCTCTCATCCCTCCCTGGGGAGCTTCTGGACAGGTCCTCTC 384
Qy 361 CTCCTTGGGCGCTCGAGAGCTCTCTGGAAACCCAGCTTCTCCACAGGCGAGGACCACA 420
Db 385 CTCCTTGGGCGCTCGAGAGCTCTCTGGAAACCCAGCTTCTCCACAGGCGAGGACCACA 444
Qy 421 GCTCACAGGATCCCAATGCCATCTCTCTGAGCTTCCCAACACCTCTCCGAGAGAGG 480
Db 445 GCTCACAGGATCCCAATGCCATCTCTCTGAGCTTCCCAACACCTCTCCGAGAGAGG 504
Qy 481 CGTTTCCCTGATCTGTAGGAGGTCACACCTCTGGTACGGCGGCGGCCAAC 533
Db 505 CGTTTCCCTGATCTGTAGGAGGTCACACCTCTGGTACGGCGGCGGCCAAC 557

RESULT 9
E12214
LOCUS Human cDNA encoding thrombopoietin. PAT 24-JUN-1998
DEFINITION E12214
ACCESSION E12214.1 GI:3251048
VERSION JP 1996291196-A/1.
KEYWORDS Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1267)
AUTHORS Miyazaki,H., Kato,H., Ogami,K., Iwamatsu,A., Akahori,H., Kuroki,R.,
Shimizu,T. and Muto,T.
TITLE PROTEIN HAVING TPO ACTIVITY
JOURNAL Patent: JP 1996291196-A 1 05-NOV-1996;
KIRIN BREWERY CO LTD
OS Homo sapiens (human)
PN JP 1996291196-A/1
PD 05-NOV-1996
PF 14-FEB-1995 JP 1995355052
PR 14-FEB-1994 JP 94P 39090, 25-MAR-1994 JP 94P 79842, PR
01-JUN-1994 JP 94P 155126, 15-JUN-1994 JP 94P 167328, PR
17-AUG-1994 JP 94P 227159, 17-AUG-1994 JP 94P 193169, PR
01-NOV-1994 JP 94P 304167, 01-DEC-1994 JP 94P 298669, PR
28-DEC-1994 JP 94P 341200
PI MIYAZAKI HIROSHI, KATO HISASHI, OGAMI KINYA, IWAMATSU AKIHIKO,
PI AKAHORI HIROAKI, KUROKI RYOTA, SHIMIZU TOSHIYUKI, PI MUTO
TAKANORI
PC C07K14/53,C07K1/22,C12N15/09//A61K38/00,C12P21/08; CC
Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..1267
FT CDS /organism='Homo sapiens'
FT /tissue_type='liver'
FT 25..1086
FT sig_peptide 25..87
FT mat_peptide 88..1083
FT /product='thrombopoietin'
FT /product='thrombopoietin'.
FEATURES source
1..1267
/organism='Homo sapiens'
/db_xref='taxon:9606'
BASE COUNT 281 a 436 c 279 g 271 t
ORIGIN

Query Match 50.3%; Score 526.6; DB 6; Length 1267;
Best Local Similarity 99.2%; Pred. No. 1.3e-99;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGGAGCTGACTGAATTCCTCGTGGTCATGCTTCTCTCACTGCAAGGCTACGCTG 60
Db 25 ATGGAGCTGACTGAATTCCTCGTGGTCATGCTTCTCTCACTGCAAGGCTACGCTG 84

```

```

Qy 61 TCCAGCCGGCTCCTCTCTGCTGTGACCTCCGAGTCCCTCAGTAACTGCTTCTGTA 120
Db 85 TCCAGCCGGCTCCTCTCTGCTGTGACCTCCGAGTCCCTCAGTAACTGCTTCTGTA 144
Qy 121 CATGTCTTTCACAGCAGACTGAGCCAGTCCGAGAGGTTACCCCTTGCCTACACCTGTC 180
Db 145 CATGTCTTTCACAGCAGACTGAGCCAGTCCGAGAGGTTACCCCTTGCCTACACCTGTC 204
Qy 181 CTGCTCCCTGCTGTGGACTTTAGCTTGGGAGATGGAACCCAGATGGAGAGACCAAG 240
Db 205 CTGCTCCCTGCTGTGGACTTTAGCTTGGGAGATGGAACCCAGATGGAGAGACCAAG 264
Qy 241 GCACAGGACATCTCTGGGAGCAGTGCACCTTCTGCTGGAGGAGTGATGGCAGCACGGGA 300
Db 265 GCACAGGACATCTCTGGGAGCAGTGCACCTTCTGCTGGAGGAGTGATGGCAGCACGGGA 324
Qy 301 CAACTGGGACCCACTTCCCTCTCATCCTCTCTGGGAGAGCTTCTGACAGAGGAGGAGTG 360
Db 325 CAACTGGGACCCACTTCCCTCTCATCCTCTCTGGGAGAGCTTCTGACAGAGGAGGAGTG 384
Qy 361 CTCTTGGGGCCCTGACAGAGCTTCTTGAACCCAGCTTCTCTCCACAGGAGGAGGAGTG 420
Db 385 CTCTTGGGGCCCTGACAGAGCTTCTTGAACCCAGCTTCTCTCCACAGGAGGAGGAGTG 444
Qy 421 GCTCACAGGATCCCAATGCCATCTCTGAGCTTCCCAACACCTGCTCCGAGGAGGAGTG 480
Db 445 GCTCACAGGATCCCAATGCCATCTCTGAGCTTCCCAACACCTGCTCCGAGGAGGAGTG 504
Qy 481 CGTTTCCCTGATCTGTAGGAGGTCACACCTCTGCTGACGGCGGCGGCCAAC 533
Db 505 CGTTTCCCTGATCTGTAGGAGGTCACACCTCTGCTGACGGCGGCGGCCAAC 557

RESULT 10
AR008878
LOCUS AR008878 1342 bp DNA
DEFINITION Sequence 1 from patent US 5756083.
ACCESSION AR008878
VERSION AR008878.1 GI:3967683
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1342)
AUTHORS Elliott,S.G.
TITLE Mpl ligand analogs
JOURNAL Patent: US 5756083-A 1 26-MAY-1998;
FEATURES Location/Qualifiers
1..1342
source
BASE COUNT 299 a 454 c 293 g 296 t
ORIGIN

Query Match 50.3%; Score 526.6; DB 6; Length 1342;
Best Local Similarity 99.2%; Pred. No. 1.3e-99;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGGAGCTGACTGAATTCCTCGTGGTCATGCTTCTCTCACTGCAAGGCTACGCTG 60
Db 36 ATGGAGCTGACTGAATTCCTCGTGGTCATGCTTCTCTCACTGCAAGGCTACGCTG 95
Qy 61 TCCAGCCGGCTCCTCTCTGCTGTGACCTCCGAGTCCCTCAGTAACTGCTTCTGTA 120
Db 96 TCCAGCCGGCTCCTCTCTGCTGTGACCTCCGAGTCCCTCAGTAACTGCTTCTGTA 155
Qy 121 CATGTCTTTCACAGCAGACTGAGCCAGTCCGAGAGGTTACCCCTTGCCTACACCTGTC 180
Db 156 CATGTCTTTCACAGCAGACTGAGCCAGTCCGAGAGGTTACCCCTTGCCTACACCTGTC 215
Qy 181 CTGCTCCCTGCTGTGGACTTTAGCTTGGGAGATGGAACCCAGATGGAGAGACCAAG 240
Db 216 CTGCTCCCTGCTGTGGACTTTAGCTTGGGAGATGGAACCCAGATGGAGAGACCAAG 275

```

Qy	241	GCACAGGACATTCGGGAGCAGTGACCCTTCTCGTGGAGGAGTGATGGCAGCACGGGA	300
Db	276	GCACAGGACATTCGGGAGCAGTGACCCTTCTCGTGGAGGAGTGATGGCAGCACGGGA	335
Qy	301	CAACTGGGACCCACTTGCCTCTCATCCCTCTGGGGCAGCTTTCTTGGACAGGTCCGCTCTC	360
Db	336	CAACTGGGACCCACTTGCCTCTCATCCCTCTGGGGCAGCTTTCTTGGACAGGTCCGCTCTC	395
Qy	361	CTCCTTGGGGCCCTGCAGAGCCCTCTTGGAAACCAGCTTCTCCACAGGGCAGGACCACA	420
Db	396	CTCCTTGGGGCCCTGCAGAGCCCTCTTGGAAACCAGCTTCTCCACAGGGCAGGACCACA	455
Qy	421	GCTCACAAAGGATCCCATAATGCCATCTTCTGAGCTTCCAACACCTGCTCCGAGGAAAGGTG	480
Db	456	GCTCACAAAGGATCCCATAATGCCATCTTCTGAGCTTCCAACACCTGCTCCGAGGAAAGGTG	515
Qy	481	CGTTTCTGTAGTCTTAGAGAGGGTCCACCCCTCTGCGTAGGGGGCGGCCAAC	533
Db	516	CGTTTCTGTAGTCTTAGAGAGGGTCCACCCCTCTGCGTAGGGGGCGGCCAAC	568
RESULT	11		
LOCUS	AR023468		
DEFINITION	Sequence 24 from patent US 5795569.	PAT	05-DEC-1998
ACCESSION	AR023468		
VERSION	AR023468.1	GI:3976762	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 1342)		
AUTHORS	Bartley,T.D., Boegenberger,J.M., Bosselman,R.A., Hunt,P., Kinsler,O.B. and Samal,B.B.		
TITLE	Mono-peylated proteins that stimulate megakaryocyte growth and differentiation		
JOURNAL	Patent: US 5795569-A 24 18-AUG-1998;		
FEATURES	Location/Qualifiers		
source	1..1342		
	/organism="unknown"		
BASE COUNT	299 a 454 c 293 g		296 t
ORIGIN			

Query Match	50.3%	Score	526.6	DB	6	Length	1342		
Best Local Similarity	99.2%	Pred. No.	1.3e-99						
Matches	529	Conservative	0	Mismatches	4	Indels	0	Gaps	0
QY	1	ATGGAGCTGACTGAATTGCTCTCCTCGTGGTCTATGCTTTCTCTTAACATGC AAGGCTTAACGGTGTG	60						
Db	36	ATGGAGCTGACTGAATTGCTCTCCTCGTGGTCTATGCTTTCTCTTAACATGC AAGGCTTAACGGTGTG	95						
QY	61	TCCAGCCCGGCTCTCTGCTTGTGACCTCGGAGTGCCTCAGTAAACTGCTTCGTGATGCC	120						
Db	96	TCCAGCCCGGCTCTCTGCTTGTGACCTCGGAGTGCCTCAGTAAACTGCTTCGTGATGCC	155						
QY	121	CATGTCCTTACAGCAGACTGAGCCAGTGGCCAGAGGTTACACCTTTTGCCCTACACCTGTC	180						
Db	156	CATGTCCTTACAGCAGACTGAGCCAGTGGCCAGAGGTTACACCTTTTGCCCTACACCTGTC	215						
QY	181	CTGCTGCCTGCTGTGGACTTTAGCTTGGGAGAAATGGAAAACCCAGATGGAGGACCAAG	240						
Db	216	CTGCTGCCTGCTGTGGACTTTAGCTTGGGAGAAATGGAAAACCCAGATGGAGGACCAAG	275						
QY	241	GCACAGGACATTCCTGGGAGCAGTGCACCTTCTGCTGGAGGGAGTGATGGCAGCACGGGGA	300						
Db	276	GCACAGGACATTCCTGGGAGCAGTGCACCTTCTGCTGGAGGGAGTGATGGCAGCACGGGGA	335						
QY	301	CAACTGGGACCACTTGCCTCTCAATCCCTCTCGGGGAGCTTTCTGTGACAGGTC CGGTCTC	360						
Db	336	CAACTGGGACCACTTGCCTCTCAATCCCTCTCGGGGAGCTTTCTGTGACAGGTC CGGTCTC	395						

Qy	361	CTCCTTGGGGCCCTGCAGAGCCTCTTGTGGAACCCAGGCTTCTCCACAGGCGAGGACCACA	420			
Db	396	CTCCTTGGGGCCCTGCAGAGCCTCTTGTGGAACCCAGGCTTCTCCACAGGCGAGGACCACA	455			
Qy	421	GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTTGCTCCGAGGAAAGGTG	480			
Db	456	GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTTGCTCCGAGGAAAGGTG	515			
Qy	481	CGTTTCTCTGATGCTTGTAGGAGGTCACACCTCTGCTCAGCGGGCGGCCAAC	533			
Db	516	CGTTTCTCTGATGCTTGTAGGAGGTCACACCTCTGCTCAGCGGGCGGCCAAC	568			
RESULT 12						
AR087573						
LOCUS	AR087573	1342 bp	DNA			
DEFINITION	Sequence 1 from patent US 5989538.					
ACCESSION	AR087573					
VERSION	AR087573.1 GI:10014336					
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 1342)					
AUTHORS	Elliott,S.G.					
TITLE	Mpl ligand analogs					
JOURNAL	Patent: US 5989538-A 1 23-NOV-1999;					
FEATURES	Location/Qualifiers					
source	1..1342					
BASE COUNT	299 a	454 c	296 t			
ORIGIN	/organism="unknown"					
Query Match 50.3%; Score 526.6; DB 6; Length 1342;						
Best Local Similarity 99.2%; Pred. No. 1.3e-99;						
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps						

Qy	1	ATGGAGCTGACTGAATTGCTCCTCGTGGTCATGCTTCTCTAACTGCAGGCTAACGGCTG	60
Db	36	ATGGAGCTGACTGAATTGCTCCTCGTGGTCATGCTTCTCTAACTGCAGGCTAACGGCTG	95
Qy	61	TCCAGCCGGCTCCTCCTGCTTGTGACCTCCGAGTCCCTCAGTAAACTGCTTCGTGACTCC	120
Db	96	TCCAGCCGGCTCCTCCTGCTTGTGACCTCCGAGTCCCTCAGTAAACTGCTTCGTGACTCC	155
Qy	121	CATGTCTTTCACAGCAGACTGAGCCAGTGCGCCAGAGGTTTCACCCTTTGCGCTTACACCTGTCT	180
Db	156	CATGTCTTTCACAGCAGACTGAGCCAGTGCGCCAGAGGTTTCACCCTTTGCGCTTACACCTGTCT	215
Qy	181	CTGCTGCTGCTGTGGACTTTTAGCTTGGGAAATGGAAACCAGATGGAGGAGACCAAG	240
Db	216	CTGCTGCTGCTGTGGACTTTTAGCTTGGGAAATGGAAACCAGATGGAGGAGACCAAG	275
Qy	241	GCACAGGACATTTCTGGGAGCAGTAGCCCTTCTGCTGGAGGAGTGATGGCAGCAGCGGGA	300
Db	276	GCACAGGACATTTCTGGGAGCAGTAGCCCTTCTGCTGGAGGAGTGATGGCAGCAGCGGGA	335
Qy	301	CAACTGGGACCACTTGGCTCTCATCCCTCTCTGGGAGAGCTTCTGGACAGGTCGGTCTC	360
Db	336	CAACTGGGACCACTTGGCTCTCATCCCTCTCTGGGAGAGCTTCTGGACAGGTCGGTCTC	395
Qy	361	CTCCTTTGGGGCCCTGCAGAGCCTCCTTGGAAACCCAGGTTCTCTCACAGGGCAGGACCA	420
Db	396	CTCCTTTGGGGCCCTGCAGAGCCTCCTTGGAAACCCAGGTTCTCTCACAGGGCAGGACCA	455
Qy	421	GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCCAACACCTGCTCCGAGGAAAGGTG	480
Db	456	GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCCAACACCTGCTCCGAGGAAAGGTG	515
Qy	481	CGTTTCCTGATGCTTGTAGGAGGTCACCCCTCTGCGTACGGCGGGCGGCCAAC	533
Db	516	CGTTTCCTGATGCTTGTAGGAGGTCACCCCTCTGCGTACGGCGGGCGGCCAAC	568

Qy	241	GCACAGGACATTC	GGAGCAGTGA	ACCCTTCTCTGCTGGAGGAGTGTATGCACGACGGGA	300					
Db	276	GCACAGGACATTC	TGGAGCAGTGA	ACCCTTCTCTGCTGGAGGAGTGTATGCACGACGGGA	335					
Qy	301	CAACTGGACCCACT	TTCGCTTCATCCCTCTGGGCAGCTTTC	TTCGGACAGGTCCGCTCTC	360					
Db	336	CAACTGGACCCACT	TTCGCTTCATCCCTCTGGGCAGCTTTC	TTCGGACAGGTTCGCTCTC	395					
Qy	361	CTCCTTGGGGCCCT	GCAGAGCCCTTGGAAACACAGTTCCTCCACAGGGCAGGACCACA	420						
Db	396	CTCCTTGGGGCCCT	GCAGAGCCCTTGGAAACACAGTTCCTCCACAGGGCAGGACCACA	455						
Qy	421	GCTCACAAAGATCC	CAATGCCATCTTCCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG	480						
Db	456	GCTCACAAAGATCC	CAATGCCATCTTCCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG	515						
Qy	481	CGTTTCTCTGATGT	TGTAGGAGGTCACCCCTCTGCTGAGGGGGGCCCAAC	533						
Db	516	CGTTTCTCTGATGT	TGTAGGAGGTCACCCCTCTGCTGAGGGGGGCCCAAC	568						
RESULT 11										
AR023468										
LOCUS	AR023468	1342 bp	DNA	PAT	05-DEC-1998					
DEFINITION	Sequence 24 from patent US 5795569.									
ACCESSION	AR023468									
VERSION	AR023468.1 GI:3976762									
KEYWORDS	Unknown.									
SOURCE	Unknown.									
ORGANISM	Unclassified.									
REFERENCE	1 (bases 1 to 1342)									
AUTHORS	Bartley,T.D., Hogenberger,J.M., Bosselman,R.A., Hunt,P., Kinstler,O.B. and Samal,B.B.									
TITLE	Mono-peylated proteins that stimulate megakaryocyte growth and differentiation									
JOURNAL	Patent: US 5795569-A 24 18-AUG-1998;									
FEATURES	Location/Qualifiers									
source	1..1342									
BASE COUNT	299 a	454 c	293 g	296 t						
ORIGIN	/organism="unknown"									
Query Match 50.3%; Score 526.6; DB 6; Length 1342;										
Best Local Similarity 99.2%; Pred. No. 1.3e-99;										
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps										
Qy	1	ATGAGCTGACTCA	ATTGCTTCCTCGTGGTCATGCTTCTCCTAACTCAAGGCTAAGCGTG	60						
Db	36	ATGAGCTGACTCA	ATTGCTTCCTCGTGGTCATGCTTCTCCTAACTCAAGGCTAAGCGTG	95						
Qy	61	TCGAGCCCGGCTC	CTCTGCTTGTGACCTCCGAGTCCCTCAGTAAACTGCTTCGTGACTCC	120						
Db	96	TCGAGCCCGGCTC	CTCTGCTTGTGACCTCCGAGTCCCTCAGTAAACTGCTTCGTGACTCC	155						
Qy	121	CATGTCTTCACAG	ACAGTGAACCCAGATGCGCAGAGTTACCCCTTGCTTACACTGTTC	180						
Db	156	CATGTCTTCACAG	ACAGTGAACCCAGATGCGCAGAGTTACCCCTTGCTTACACTGTTC	215						
Qy	181	CTGCTGCCTCTG	TGACCTTACCTTGGGAGATGGAACCCAGATGAGGAGACCAAG	240						
Db	216	CTGCTGCCTCTG	TGACCTTACCTTGGGAGATGGAACCCAGATGAGGAGACCAAG	275						
Qy	241	GCACAGGACATTC	TGGAGCAGTGA	ACCCTTCTCTGCTGGAGGAGTGTATGCACGACGGGA	300					
Db	276	GCACAGGACATTC	TGGAGCAGTGA	ACCCTTCTCTGCTGGAGGAGTGTATGCACGACGGGA	335					
Qy	301	CAACTGGACCCACT	TTCGCTTCATCCCTCTGGGCAGCTTTC	TTCGGACAGGTCCGCTCTC	360					
Db	336	CAACTGGACCCACT	TTCGCTTCATCCCTCTGGGCAGCTTTC	TTCGGACAGGTCCGCTCTC	395					

RESULT 13
185129
LOCUS 185129 1342 bp DNA PAT 04-APR-1998
DEFINITION Sequence 1 from patent US 5696250.
ACCESSION 185129
VERSION 185129.1 GI:3022649
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 1342)
AUTHORS Elliott,S.G.
TITLE DNA encoding megakaryocyte growth and development factor analogs
JOURNAL Patent: US 5696250-A 1 09-DEC-1997;
FEATURES
Location/Qualifiers
source
1..1342
/organism="unknown"
BASE COUNT 299 a 454 c 293 g 296 t
ORIGIN

Query Match 50.3%; Score 526.6; DB 6; Length 1342;
Best Local Similarity 99.2%; Pred. No. 1.3e-99;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATTCCTCGTGGTTCATGCTTCTCTTAAGTCAAGGCTAAGCGTG 60
Db
36 ATGGAGCTGACTGAATTCCTCGTGGTTCATGCTTCTCTTAAGTCAAGGCTAAGCGTG 95
QY 61 TCCAGCCCGGCTCTCTCTGCTGTGACCTCCGAGTCCCTCAGTAACTGCTTGGTACTCC 120
Db
96 TCCAGCCCGGCTCTCTCTGCTGTGACCTCCGAGTCCCTCAGTAACTGCTTGGTACTCC 155
QY 121 CATGCTCTCACAGACAGTGAAGCCAGTCCGAGAGGTTTCAACCTTGTGCTACACTGTC 180
Db
156 CATGCTCTCACAGACAGTGAAGCCAGTCCGAGAGGTTTCAACCTTGTGCTACACTGTC 215
QY 181 CTGCTGCTCTGCTGACATTTAGCTTGGAGAGTGAAGAACCCAGATGGAGAGACCAAG 240
Db
216 CTGCTGCTCTGCTGACATTTAGCTTGGAGAGTGAAGAACCCAGATGGAGAGACCAAG 275
QY 241 GCACAGGACATTTGGAGACAGTGAACCTTCTGCTGGAGGAGTGAAGGAGGAGGGA 300
Db
276 GCACAGGACATTTGGAGACAGTGAACCTTCTGCTGGAGGAGTGAAGGAGGAGGGA 335
QY 301 CAACTGGGACCACTTGCCTCTCATCCCTCTGCTGGGACAGCTTCTGGACAGTCCGTC 360
Db
336 CAACTGGGACCACTTGCCTCTCATCCCTCTGCTGGGACAGCTTCTGGACAGTCCGTC 395
QY 361 CTCTTTGGGACCTCGAGAGCCTTCTTGAACCCAGCTTCTTCCACAGGACGACCA 420
Db
396 CTCTTTGGGACCTCGAGAGCCTTCTTGAACCCAGCTTCTTCCACAGGACGACCA 455
QY 421 GCTCACAAGGATCCCAATGCCATTCCTTCCTGAGCTTCCACACCTGCTCCGAGGAAAGTG 480
Db
456 GCTCACAAGGATCCCAATGCCATTCCTTCCTGAGCTTCCACACCTGCTCCGAGGAAAGTG 515
QY 481 CGTTTCTGATGCTTAGAGAGGTTCCACCTCTCGGTACGGCGGGCGCCCAAC 533
Db
516 CGTTTCTGATGCTTAGAGAGGTTCCACCTCTCGGTACGGCGGGCGCCCAAC 568

RESULT 14
D32047
LOCUS D32047 1691 bp mRNA PRI 03-SEP-1997
DEFINITION Homo sapiens mRNA for thrombopoietin, complete cds.
ACCESSION D32047
VERSION D32047.1 GI:2351117
KEYWORDS
SOURCE Homo sapiens liver cdna to mRNA.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1691)
Ogami, K.
Direct Submission
Submitted (08-JUL-1994) to the DDBJ/EMBL/GenBank databases. Kinya
Ogami, KIRIN Brewery Co., Ltd., Pharmaceutical Research Laboratory;
2-2 Soujamachi 1 chome, Maebashi, Gunma 371, Japan
(Tel:0272-54-8618, Fax:0272-52-2307)
2 (sites)
Kato, T., Ogami, K., Shimada, Y., Iwamatsu, A., Sohma, Y., Akahori, H.,
Horie, K., Kokubo, A., Kudo, Y., Maeda, E., Kobayashi, K., Onashi, H.,
Ozawa, T., Inoue, H., Kawamura, K. and Miyazaki, H.
Purification and characterization of thrombopoietin
J. Biochem. 118 (1), 229-236 (1995)
96015174
Erratum: [[published erratum appears in J Biochem (Tokyo) 1996
Jan;119(1):208]]
Location/Qualifiers
1..1691
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="liver"
102..1163
102..1163
/codon_start=1
/product="thrombopoietin"
/protein_id="BAA21930.1"
/db_xref="GI:2351118"
/translation="MELTELLLVMLLTARLTLSPPAPCDRLVSLKLLRSHVLH
SRLSQCEVHELPTVLLPAVDLSLGEMKTOEETKAQDILGAVTLLEGVMAARGQL
PTCLSLGSLGSLGQVRLGALGSLGTLQPPQRTTAHKQNPATLFSQHLRLGKV
RFLMVGSTLCLVRRAPPTAVPSRTSLVLTNLELNPRTSGLETFASARTLGGSL
LKWQGFRAKIPGLLNOTSRSLDOIPGYLNRHELLNGTRGLFPGSRRTLGAPDLSL
GTSDTGLPPNLPQGYSPSPHPPTGQVTLFPLPTLTPVVLHLLPLDPSAPTTP
TSPLLNTSYTHSQNLSEBQ"
BASE COUNT 383 a 542 c 371 g 395 t
ORIGIN

Query Match 50.3%; Score 526.6; DB 9; Length 1691;
Best Local Similarity 99.2%; Pred. No. 1.2e-99;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATTCCTCGTGGTTCATGCTTCTCTTAAGTCAAGGCTAAGCGTG 60
Db
102 ATGGAGCTGACTGAATTCCTCGTGGTTCATGCTTCTCTTAAGTCAAGGCTAAGCGTG 161
QY 61 TCCAGCCCGGCTCTCTCTGCTGTGACCTCCGAGTCCCTCAGTAACTGCTTGTGACTCC 120
Db
162 TCCAGCCCGGCTCTCTCTGCTGTGACCTCCGAGTCCCTCAGTAACTGCTTGTGACTCC 221
QY 121 CATGCTCTCACAGCAGACTGAGCCAGTCCGAGAGTTCACCTTTGGCTACACCTGTC 180
Db
222 CATGCTCTCACAGCAGACTGAGCCAGTCCGAGAGTTCACCTTTGGCTACACCTGTC 281
QY 181 CTGCTGCTCTGCTGGACTTTAGCTTGGAGAGTGAAGAACCCAGATGGAGGAGACCAAG 240
Db
282 CTGCTGCTCTGCTGGACTTTAGCTTGGAGAGTGAAGAACCCAGATGGAGGAGACCAAG 341
QY 241 GCACAGGACATTTGGAGCAGTGAACCTTCTTGAACCCAGCTTCTCCACAGGACGACCA 300
Db
342 GCACAGGACATTTGGAGCAGTGAACCTTCTTGAACCCAGCTTCTCCACAGGACGACCA 401
QY 301 CAACCTGGGACCCACTTGGCTCTCATCCCTCTGGGACAGCTTCTGCGACAGTCCGTC 360
Db
402 CAACCTGGGACCCACTTGGCTCTCATCCCTCTGGGACAGCTTCTGCGACAGTCCGTC 461
QY 361 CTCTCTGGGACCCCTCGACAGCCTCTTGAACCCAGCTTCTCCACAGGACGACCA 420
Db
462 CTCTCTGGGACCCCTCGACAGCCTCTTGAACCCAGCTTCTCCACAGGACGACCA 521
QY 421 GCTCACAAGGATCCCAATGCCATTCCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGTG 480

|||||
Db 522 GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCAACACCTGCTCCGAGGAAAGGTG 581
|||||
Qy 481 CGTTTCTGTAGCTGTAGGAGGTCCACCTCTGCGTACGGCGGCCCAAC 533
|||||
Db 582 CGTTTCTGTAGCTGTAGGAGGTCCACCTCTGCGTACGGCGGCCCAAC 634
|||||
RESULT 15
E16668
LOCUS E16668 1721 bp DNA PAT 28-JUL-1999
DEFINITION cDNA encoding thrombopoietin.
ACCESSION E16668
VERSION E16668.1 GI:57111351
KEYWORDS JP 1998212243-A/2.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1721)
AUTHORS Kuroda K
TITLE COMPOSITION CONTAINING THROMBOPOIETIN PROTEIN AND USED FOR
JOURNAL STIMULATING RECOVERY OF HEMOSTATIC THROMBUS-FORMING ABILITY OF
PATENT: JP 1998212243-A 2 11-AUG-1998;
KIRIN BREWERY CO LTD
COMMENT OS Homo sapiens (human)
PN JP 1998212243-A/2
PD 11-AUG-1998
PF 31-JAN-1997 JP 1997019549
PI KURODA KENJI
PC A61K38/00,A01M1/02,C07K14/52//C12N15/09,C12P21/02,(C12P21/02,
C12R1:19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FH 1..1721
FT source /organism='Homo sapiens'
FT /tissue_type='Liver'
FT 1..101
FT 5'UTR
FT sig_peptide 102..164
FT /product='signal peptide of thrombopoietin' FT
mat_peptide 165..1160
FT /product='thrombopoietin'
FT CDS 102..1163
FT /product='thrombopoietin'
FT 1164..1721.
FT 3'UTR Location/Qualifiers
source 1..1721
/organism='Homo sapiens'
/db_xref='taxon:9606'
BASE COUNT 413 a 542 c 371 g 395 t
ORIGIN

Query Match 50.3%; Score 526.6; DB 6; Length 1721;
Best Local Similarity 99.2%; Pred. No. 1.2e-99;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 ATGAGCTGACTGAATGCTCCGTGGTATGCTTCTTCTTAAGTCAAGGCTAACGCTG 60
|||||
Db 102 ATGAGCTGACTGAATGCTCCGTGGTATGCTTCTTCTTAAGTCAAGGCTAACGCTG 161
|||||
Qy 61 TCAGCCCGGCTCCTCTGCTGTGACCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
|||||
Db 162 TCAGCCCGGCTCCTCTGCTGTGACCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 221
|||||
Qy 121 CATGTCCTTCACAGCAGCTGAGCCAGTGCACAGGTTTCAACCTTTGCGTTACACCTGTC 180
|||||
Db 222 CATGTCCTTCACAGCAGCTGAGCCAGTGCACAGGTTTCAACCTTTGCGTTACACCTGTC 281
|||||

Qy 181 CTGCTGCTGCTGTGAGCTTTAGCTTGGGAGATGGAACCCAGATGGAGGAGACCAAG 240
|||||
Db 282 CTGCTGCTGCTGTGAGCTTTAGCTTGGGAGATGGAACCCAGATGGAGGAGACCAAG 341
|||||
Qy 241 GCACAGGACATTCCTGGGAGCAGTACCCCTTCTGCTGAGGGAGTGTGGCAGCAGGGGA 300
|||||
Db 342 GCACAGGACATTCCTGGGAGCAGTACCCCTTCTGCTGAGGGAGTGTGGCAGCAGGGGA 401
|||||
Qy 301 CAACTGGGACCCACTTGCCTCTCATCCCTCTGGGGCAGCTTTCTGGACAGGTCCGTCTC 360
|||||
Db 402 CAACTGGGACCCACTTGCCTCTCATCCCTCTGGGGCAGCTTTCTGGACAGGTCCGTCTC 461
|||||
Qy 361 CTCCTTGGGGCCCTGCAGAGCCTCCTTTGGAAACCCAGCTTCTCCACAGGGCAGGACACA 420
|||||
Db 462 CTCCTTGGGGCCCTGCAGAGCCTCCTTTGGAAACCCAGCTTCTCCACAGGGCAGGACACA 521
|||||
Qy 421 GCTCACAAAGATCCCAATGCCATCTTCTGAGCTTCCAACTGCTCCGAGGAAAGGTG 480
|||||
Db 522 GCTCACAAAGATCCCAATGCCATCTTCTGAGCTTCCAACTGCTCCGAGGAAAGGTG 581
|||||
Qy 481 CGTTTCTGTAGCTGTAGGAGGTCCACCTCTGCGTACGGCGGGGCCCAAC 533
|||||
Db 582 CGTTTCTGTAGCTGTAGGAGGTCCACCTCTGCGTACGGCGGGGCCCAAC 634
|||||

Search completed: April 5, 2002, 10:29:29
Job time: 4762 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2002, 09:10:41 ; Search time 188.75 Seconds
(without alignments)
4755.600 Million cell updates/sec

Title: US-09-680-514-4

Perfect score: 1047

Sequence: 1 ATGGAGCTGACTGAATTGCT.....TACGCCACCTTGCCAGGCC 1047

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_1101.*

- 1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.*
- 2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1047	100.0	1047	17	AA141786
2	993	94.8	1093	17	AA141787
3	987.4	94.3	1095	17	AA141788
4	526.6	50.3	861	16	AAQ99552
5	526.6	50.3	1062	16	AA103941
6	526.6	50.3	1062	16	AA104050
7	526.6	50.3	1062	17	AA134852
8	526.6	50.3	1062	17	AA137383
9	526.6	50.3	1062	17	AA132591
10	526.6	50.3	1062	18	AA185555
11	526.6	50.3	1062	19	AA121696

12	526.6	50.3	1062	21	AAA51991	Human thrombopoietin
13	526.6	50.3	1086	17	AA147958	Human thrombopoietin
14	526.6	50.3	1342	16	AAQ99704	Human MGDF-1 and M
15	526.6	50.3	1342	17	AA136657	Native human mpl 1
16	526.6	50.3	1342	17	AA110025	Human MGDF-1/2 CDN
17	526.6	50.3	1721	16	AAQ99554	Thrombopoietin cod
18	526.6	50.3	1721	21	AA240191	Human wild type th
19	526.6	50.3	1774	21	AA447793	Human thrombopoiet
20	526.6	50.3	1795	16	AAQ94107	hML cDNA. Homo sa
21	526.6	50.3	1795	18	AA164318	Human thrombopoiet
22	525	50.1	600	17	AA136658	Truncated human mp
23	525	50.1	605	20	AA132813	Human truncated mp
24	525	50.1	1267	16	AAQ99553	Thrombopoietin cod
25	514	49.1	525	11	AAQ04482	Plasmid PASN6 enco
26	510.8	48.8	525	11	AAQ04484	Plasmid PASN145 en
27	506	48.3	525	11	AAQ04481	Plasmid PAS28 enco
28	503.6	48.1	1342	17	AA133933	Human megakaryocyt
29	503.6	48.1	1342	19	AA129068	Nucleotide sequenc
30	503.6	48.1	1342	20	AA132812	Human mpl ligand e
31	492.6	47.0	1050	17	AA141901	Human thrombopoiet
32	486.8	46.5	525	9	AA180947	G-CSF gene isolate
33	486.8	46.5	1520	8	AA170223	Plasmid pBRV2 inse
34	486.8	46.5	1520	8	AA171320	Sequence encoding
35	486.8	46.5	1521	7	AA181478	Plasmid pBRV2 inse
36	486.8	46.5	1525	9	AA181478	Sequence encoding
37	486.8	46.5	1525	10	AA191086	Plasmid pP12 contg
38	485.2	46.3	1415	8	AA171089	Sequence of human
39	479.8	45.8	741	17	AA141907	Human thrombopoiet
40	478	45.7	861	17	AA141902	Human thrombopoiet
41	478	45.7	861	17	AA15486	Megakaryocyte diff
42	478	45.7	946	17	AA129735	Human platelet pro
43	478	45.7	946	21	AA244174	Human platelet-pro
44	478	45.7	946	21	AA244179	Human platelet-pro
45	478	45.7	1164	16	AAQ99705	Human MGDF-3. Hom

ALIGNMENTS

RESULT 1
AA141786
ID AAT41786 standard; DNA; 1047 BP.
XX
AC AAT41786;
XX
DT 01-JUL-1997 (first entry)
XX
DE Fusion peptide #1 having G-CSF and TPO activity.
XX
KW Fusion protein; human granulocyte colony stimulating factor; hG-CSF;
KW thrombopoietin; TPO; spacer peptide; blood platelet production;
KW leukocyte production; anaemia; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT sig_peptide 1..63
FT mat_peptide 64..1047
FT /*tag= a
FT /*tag= b
XX
XX WO9634016-A1.
XX
PD 31-OCT-1996.
XX
PF 26-APR-1996; 96WO-JP01157.
XX
PR 26-APR-1995; 95JP-0102625.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Anazawa H, Konishi N, Shiotsu Y, Tamaoki T, Terasaki Y;
PI Uchida K, Yamasaki M, Yamashita K, Yokoi H;

(TPO) activity, opt. via a spacer peptide. Peptides derived from these by deletion, insertion or substitution of one or more amino acid residues are included within the scope of the invention. The fusion peptides stimulate blood platelet and leukocyte production and are useful in the treatment of anaemia.

SQ Sequence 1083 BP; 197 A; 355 C; 305 G; 226 T; 0 other;

Query Match 94.8%; Score 993; DB 17; Length 1083;
Best Local Similarity 96.2%; Pred. NO. 1.9e-228;
Matches 1042; Conservative 0; Mismatches 5; Indels 36; Gaps 1;

Qy	1	ATGAGCTGACTGAATTCGCTCTCTCGTGGTTCATGCTTCTCTTA	60
Db	1	atgagctgactgaattgctctctggttcattgcttctcttaactgaagctaa	60
Qy	61	TCCAGCCGGCTCCTCTGCTGTGACCTCCGAGTCTCAGTAACTGCT	120
Db	61	tccagccggctcctctgctgttgacctccgagtcctcaagtaaaactgcttg	120
Qy	121	CATGTCCTTACAGCAGACTGAGCCAGTGCCCGAGAGTTACCCCTT	180
Db	121	catgtccttcacagcagactgagccagtgcccagaggttcaccttggctac	180
Qy	181	CTGCTGCTGCTGGACTTTAGCTTCGGAGATGGAARACCCAGATG	240
Db	181	ctgctgctgctggactttagcttcggagatggaaccaccagatgagggag	240
Qy	241	GCACAGGACATTCGGGAGCAGTGACCTTCTGCTGGAGGGAGTGT	300
Db	241	gcacaggacattcgggagcagtgaacctctctgctgagggagtgatggc	300
Qy	301	CAACTGGGACCACTTGCCTTCATCCCTCCTGGGCAGCTTTC	360
Db	301	caactgggaccaaacttgctctctcatcctctggtggcaggttcttgag	360
Qy	361	CTCTTGGGGCCCTGCAGAGCCTCTTGGAAACCAGCTTCTCCAC	420
Db	361	ctcttggggccctgcagagcctcttggaaccacgcttctccaaggagga	420
Qy	421	GCTCACAAAGNATCCATGCGATTCCTGAGCTTCCACACCTGCT	480
Db	421	gctcacaaagatccaatgcatctctctgaggttccaacacgtctccgag	480
Qy	481	CGTTTCTGTGCTGTGAGGAGGTCCACCTCTGCTGACGGCGG	525
Db	481	cgtttctgtatgctgttagggagggtccaccctctgctcaagggttctg	540
Qy	526	-----GGCCCAACATATCGCGCTTCGAGTCTACACAGAGCTTC	564
Db	541	ggttcggagggggttctagagcacaacacatcgcgctcgsagtctacc	600
Qy	565	CTTTTAAAAAGCTTAGAGCAAGTCAGGAATCCAGGGCATGGCC	624
Db	601	cttttaaaagcttagagcaagtgaggaagatccaaggagatggcgacg	660
Qy	625	AAGCTGTGTGCCACTACAAGCTGTGCCACCCGAGGAGCTGTGT	684
Db	661	aagctgtgtgccacctacaagctgtgccaccccgaggagctgtgtctg	720
Qy	685	CTGGGCATCCCTTGGGCTCCCTGAGGAGCTGCCACGCGCCTTG	744
Db	721	ctgggcataccctgggtcccttgagcagtgcccacgacaggccctgag	780
Qy	745	TGCTGTAGCCCAACTTCCATAGCGGCTTTTCTCTACACAGGGCT	804
Db	781	tgcttgagccaactccatagcgccctttctctaccagggtctcgaagcc	840
Qy	805	GGGATCTCCCGAGTTGGGTGCCACCTTGGACACACTGCAGCTG	864
Db	841	gggatctccccaggttggttcccacttgacacactgagctggaagctgc	900

QY	865	GCCACCACCACTCTGGCAGCAGATGGAGAACTGGGAATGGCCCTTGCCCTGCAGCCACCC	924
Db	901	gccaccaccaactctggcagcagatgggaagaactgggaatggccctggccctgcagcccaacc	960
QY	925	CAGGGTGCCATGCGCGGCTTTCGCCTCTGCTTTTCCAGCGCGGGCAGAGGGGTCTCTAGTT	984
Db	961	cagggtgccatgcgcggccttcgcctctgcttcccgccgagggggtcctcagtt	1020
QY	985	GCCTCCCATCTGCAGAGCTTCCTGGAGGTGCGTACCGCGTTCATCGCCACCTTGCCCCAG	1044
Db	1021	gcctcccatctgcagagcttctcggaggtgctgtacgcgcttctacgccaactctgcccag	1080
QY	1045	CCC 1047	
Db	1081	ccc 1083	
RESULT	3		
AAAT41788			
ID	AAAT41788	standard; DNA; 1095 BP.	
XX	AC		
XX	AAAT41788;		
DT	01-JUL-1997	(first entry)	
XX		Fusion peptide #3 having G-CSF and TPO activity.	
DE			
XX			
KW		Fusion protein; human granulocyte colony stimulating factor; hG-CSF;	
KW		thrombopoietin; TPO; spacer peptide; blood platelet production;	
KW		leukocyte production; anaemia; ds.	
XX			
OS		Homo sapiens.	
XX			
FH	Key	Location/Qualifiers	
FT	sig_peptide	1..63	
FT		/*tag= a	
FT	mat_peptide	64..1095	
FT		/*tag= b	
XX			
PN	W09634016-A1.		
XX			
PD	31-OCT-1996.		
PF	26-APR-1996;	96WO-JP01157.	
XX			
PR	26-APR-1995;	95JP-0102625.	
XX			
PA	(KYOW) KYOWA HAKKO KOGYO KK.		
XX			
PI	Anazawa H, Konishi N, Shiotsu Y, Tamaoki T, Terasaki Y;		
PI	Uchida K, Yamasaki M, Yamashita K, Yokoi H;		
XX			
DR	WPI: 1996-497573/49.		
DR	P-PSDB; AAW00379.		
XX			
PT		Fusion peptide having G-CSF and with thrombopoietin activity -	
PT		optionally chemically modified with a poly(alkylene glycol, used for	
PT		treatment of anaemia	
XX			
PS	Claim 3; Page 52-54; 71pp; Japanese.		
XX			
CC		The sequences given in AAAT41786-88 encode fusion proteins which are	
CC		composed of a peptide having human granulocyte colony stimulating	
CC		factor (hG-CSF) activity fused with a peptide having thrombopoietin	
CC		(TPO) activity, opt. via a spacer peptide. Peptides derived from	
CC		these by deletion, insertion or substitution of one or more amino	
CC		acid residues are included within the scope of the invention. The	
CC		fusion peptides stimulate blood platelet and leukocyte production	
CC		and are useful in the treatment of anaemia.	
XX			
SQ	Sequence 1095 BP; 194 A; 362 C; 311 G; 228 T; 0 other;		

Sequence 1095 BP; 194 A; 362 C; 311 G; 228 T; 0 other;

AA		
FH	Key	Location/Qualifiers
FT	CDS	1..1062
FT		/*tag= a

XX

RESULT 9
AAT32591

Db	241	gcacagacattctggagcagtgaccttctgctgagggagtgtatgagcagcacgggga	300
QY	301	CAACTGGGACCCACTTGGCTCTCATCCCTCCGCGGACAGCTTCTTGACAGTCCGTCTC	360
Db	301	caactgggaccacacttgctctctcatctctctggggcagcttcttgacaggtccgtctc	360
QY	361	CTCCTTGGGGCCCTGCAGAGCCCTCTTGGAAACCACAGCTTCCTCCACAGGGCAGGACCACA	420
Db	361	ctccttggggccctgcagagctctcttgaaacccagcttcttccacagggcaggaccaca	420
QY	421	GCTCACAAGSANTCCCAATCCGATCTCTCTGAGTTCCTCCACACCTGCTCCGAGGAAAGTG	480
Db	421	gctcacaaggatcccaatgcatctctctgagcttccaacacctgctccgaggaaggtg	480
QY	481	CGTTTCCTGATGCTTGTAGGAGGTCCACCTCTCTGCGTACGGCGGGCGCCCAAC	533
Db	481	cgtttctctgatgctttagggaggtccacctctgctcaggcgggcccccacc	533
RESULT 10			
AAT85555			
ID	AAT85555 standard; cDNA; 1062 BP.		
XX			
AC	AAT85555;		
DT	05-NOV-1997 (first entry)		
XX			
DE	Human thrombopoietin cDNA.		
XX			
KW	TPO; cancer; myeloid cell proliferation; expansion; bone marrow;		
KW	peripheral blood stem cell; chemotherapy; radiation therapy;		
KW	breast cancer; leukaemia; lymphoma; multiple myeloma; ds.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..1062	
FT		/*tag= a	
FT	mat_peptide	64..1059	
FT		/*tag= b	
FT		/product= Thrombopoietin	
XX			
PN	W09640218-A1.		
XX			
PD	19-DEC-1996.		
XX			
PF	29-MAY-1996; 96WO-US07880.		
XX			
PR	07-JUN-1995; 95US-0482212.		
XX			
PA	(ZYMO) ZYMOGENETICS INC.		
XX			
PI	Fibbe WE, Grossmann A;		
XX			
DR	WPI; 1997-108626/10.		
XX			
DR	P-PSDB; AAW26841.		
XX			
PT	Increasing haematopoietic cells in patient - by admin. of stem cells		
PT	from donor previously treated with thrombopoietin, used in patients		
PT	who have received chemotherapy or radiation for cancer		
XX			
PS	Disclosure; Page 17-19; 32pp; English.		
XX			
CC	In a claimed method, thrombopoietin (TPO) (especially human TPO) is		
CC	administered to a donor to stimulate proliferation of myeloid cells.		
CC	Bone marrow or peripheral blood stem cells are collected from the		
CC	donor and administered to the recipient. The method is used for		
CC	preparing cells for (bone marrow) transplantation and for stimulating		
CC	platelet or erythrocyte recovery in a patient receiving chemotherapy		
CC	or radiation therapy for e.g. breast cancer, leukaemia, lymphoma or		
CC	multiple myeloma. Treating the donor with TPO accelerates restoration		
CC	of erythrocyte and thrombocyte levels in the patient after transplant.		

CC This restoration may be further improved by administering TPO to the
XX recipient after transplant. The present sequence encodes human TPO.
SQ Sequence 1062 BP; 223 A; 375 C; 236 G; 228 T; 0 other;

Query Match 50.3%; Score 526.6; DB 18; Length 1062;
Best Local Similarity 99.2%; Pred. No. 7.5e-117;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATGCTCCTCGTGTGATGCTTCTCTAACTGCAAGGCTAACCGTG 60
Db 1 atggagctgactgaatgctcctcggtgtgatgcttctctaaactgcaaggctaacgctg 60
QY 61 TCCAGCCGGCTCCTCGCTGTTGACTCCGAGTCCGAGTCTCAGTAAACTGCTTCGAGACTCC 120
Db 61 tccagccggctcctcgctgttgactccgagtcctcagactcctcagtaactgcttcgactcc 120
QY 121 CATGCTCTTACAGCAGACTGAGCCAGTGGCCAGAGGTTTACCCCTTTGCCCTACACCTGTC 180
Db 121 catgctcttacagcagactgagccagtgccagaggttcaacctttgctacacctgtc 180
QY 181 CTGCTGCTGCTGTGGAGCTTTAGCTTGGGAGAAATGGAAACCCAGATGGAGGAGACCAAG 240
Db 181 ctgctgctgctgtggactttagcttgggagaaatggaaacccagatggagagaccag 240
QY 241 GCACAGGACATTCCTGGGAGAGTACACCTTCTGCTGGAGGAGTGGAGGAGACGGGA 300
Db 241 gcacaggacatctcctgggagagtgaccttctgctggagggagtgaaggcagcagggga 300
QY 301 CAATGGGACCCACTTGCCTCTCATCCCTCCTGGGGGAGGTTTCTGGACAGGTCCTGTC 360
Db 301 caatgggacccacttgcctctcatccctcctggggaggttctggacaggtccgtctc 360
QY 361 CTCTTTGGGCGCTGCGAGAGCTCCTTGGAAACCCAGCTTCTTCCACAGGGGAGACCA 420
Db 361 ctctttgggcgctgctgagagctccttggaaacccagcttcttccacagggcagaccaca 420
QY 421 GCTCACAGGATCCCAATGCATCTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGTG 480
Db 421 gctcacaggatcccaatgcctctctgagcttccaaacacctgctccgaggaaggtg 480
QY 481 CGTTCCTGATGCTTGTAGGAGGTCACCCCTCTGCTAGCGGGCGCCCAAC 533
Db 481 cgttctctgatgcttgtagagggttcacacctctgctcaggcggcccccacc 533

RESULT 11

AAV21696
AAV21696 standard; cDNA; 1062 BP.

AAV21696;

07-AUG-1998 (first entry)

Human thrombopoietin encoding cDNA.

Thrombopoietin; TPO; expression vector; platelet number; cytokine;
human; bone marrow; proliferation; treatment; cancer; aplastic anaemia;
myelodisplastic syndrome; chemotherapy; cytopenia; thrombocytopenia;
haematologic disorder; leukaemia; lymphoma; ss.

Homo sapiens.

Key	Location/Qualifiers
FT CDS	1..1062
FT	/*tag= a
FT	/product= "Thrombopoietin polypeptide"
FT	1..63
FT	/*tag= b
FT	64..1059
FT	/*tag= c

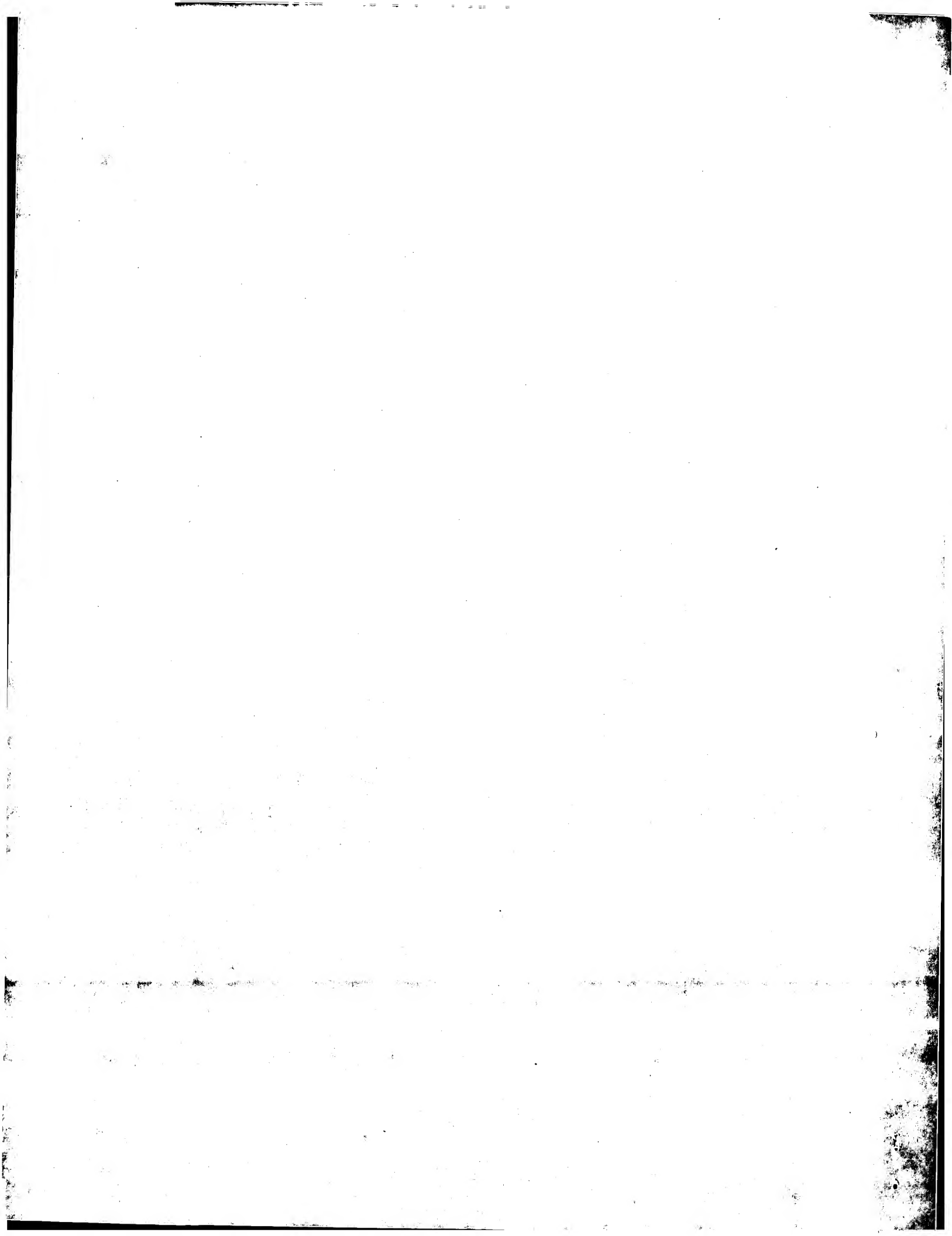
PN WO9806849-A1.
XX 19-FEB-1998.
XX 30-JUL-1997; 97WO-US13543.
XX 13-AUG-1996; 96US-0696447.
XX (ZYMO) ZYMOGENETICS INC.
XX Irani M, Morrison-nelson GR;
XX WPI; 1998-159541/14.
XX P-PSDB; AAW53124.
XX Thrombopoietin protein expression vector - used for increasing
XX platelet number in a mammal
XX Disclosure; Pages 29-32; 56pp; English.

CC This cDNA encodes a thrombopoietin (TPO) polypeptide. This can be used
in the construction of a new expression vector replicable in a
eukaryotic host cell encoding TPO polypeptides. The vector comprises a
transcription promoter, a first DNA segment encoding a secretory leader,
a second segment encoding a TPO polypeptide and a transcription
terminator which are all operably linked. The second TPO segment consists
of C-X-B, where C is a human TPO cytokine domain, X is a peptide bond or
a linker consisting of one or two amino acid residues, such that X along
in combination with C or B does not provide a dibasic amino acid pair
and B is a polypeptide that can be selected from AAW53125 to AAW53126.
The secretory leader is a S. cerevisiae alpha-factor secretory leader. A
cultured yeast cell containing such an expression vector can be used to
produce the TPO polypeptide. The TPO polypeptide can be used in a method
for increasing platelet number in a mammal. It can be used to increase
proliferation of bone marrow cells for treatment of cytopenia, including
those induced by aplastic anaemia, myelodisplastic syndromes,
chemotherapy or congenital cytopenias. It can also be used to treat
thrombocytopenia, haematologic disorders, such as leukaemia and lymphoma
or metastatic cancers involving bone marrow.

SQ Sequence 1062 BP; 223 A; 375 C; 236 G; 228 T; 0 other;

Query Match 50.3%; Score 526.6; DB 19; Length 1062;
Best Local Similarity 99.2%; Pred. No. 7.5e-117;
Matches 529; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATGCTCCTCGTGTGATGCTTCTCTAACTGCAAGGCTAACCGTG 60
Db 1 atggagctgactgaatgctcctcggtgtgatgcttctctaaactgcaaggctaacgctg 60
QY 61 TCCAGCCGGCTCCTCGCTGTTGACTCCGAGTCCGAGTCTCAGTAAACTGCTTCGAGACTCC 120
Db 61 tccagccggctcctcgctgttgactccgagtcctcagtaactgcttcgactcc 120
QY 121 CATGCTCTTACAGCAGACTGAGCCAGTGGCCAGAGGTTTACCCCTTTGCCCTACACCTGTC 180
Db 121 catgctcttacagcagactgagccagtgccagaggttcaacctttgctacacctgtc 180
QY 181 CTGCTGCTGCTGTGGAGCTTTAGCTTGGGAGAAATGGAAACCCAGATGGAGGAGACCAAG 240
Db 181 ctgctgctgctgtggactttagcttgggagaaatggaaacccagatggagagaccag 240
QY 241 GCACAGGACATTCCTGGGAGGAGTACACCTTCTGCTGGAGGAGTGGAGGAGACGGGA 300
Db 241 gcacaggacatctcctgggagagtgaccttctgctggagggagtgaaggcagcagggga 300
QY 301 CAACTGGGACCCACTTGCCTCTCATCCCTCCTGGGGGAGGTTTCTGGACAGGTCCTGTC 360
Db 301 caactgggacccacttgcctctcatccctcctctctctctctctctctctctctctctc 360
QY 361 CTCTTTGGGCGCTGCGAGAGCTCCTTGGAAACCCAGCTTCTTCCACAGGGGAGACCA 420
Db 361 ctctttgggcgctgctgagagctccttggaaacccagcttcttccacagggcagaccaca 420



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2002, 08:51:36 ; Search time 1909.33 seconds
(without alignments)
5892.554 Million cell updates/sec

Title: US-09-680-514-4
Perfect score: 1047
Sequence: 1 ATGGAGCTGACTGAATTGCT.....TACGCCACCTTCCCGAGCCC 1047

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_htc:*
10: gb_est1:*
11: gb_est2:*
12: gb_htc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pin:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	320.4	30.6	973	11	BI411128
2	312.8	29.9	598	11	BG548320
3	301.6	28.8	900	11	BI328788
4	300.6	28.7	507	10	BE485194
5	300.6	28.7	633	10	AI314551
6	299.6	28.6	609	10	AA538257
7	285.2	27.2	878	11	BF789308
8	230	22.0	464	10	BE754587
9	135.8	13.0	327	11	BF848766
10	135.2	12.9	449	13	AZ025993
11	109.4	10.4	419	13	AZ352219
12	86.4	8.3	708	13	AZ792988
13					2M0046105

13	77.8	7.4	752	13	AQ888740
14	65.6	6.3	589	10	BE377769
15	54.8	5.2	925	13	CNS0091P
16	50.2	4.8	513	10	BE480590
17	49	4.7	925	13	CNS0091P
18	48.8	4.7	617	10	AI981598
19	43.2	4.1	461	10	BE598945
20	43.2	4.1	559	11	BG556176
21	42.6	4.1	452	10	AA155632
22	42.6	4.1	899	13	CNS015V7
23	42.4	4.0	1203	13	CNS015V4
24	42.2	4.0	598	10	AA115932
25	41.4	4.0	421	10	AA058743
26	41.4	4.0	478	11	BI345642
27	41.2	3.9	1201	13	CNS0141B
28	40.8	3.9	545	11	BF868325
29	40.6	3.9	447	11	BG604563
30	40.6	3.9	845	10	AL572931
31	40.6	3.9	932	13	CNS0072Q
32	40.6	3.9	1003	11	BI410408
33	40.6	3.9	1101	13	CNS05709
34	40.4	3.9	458	10	AW273202
35	40.4	3.9	485	10	BE463718
36	40.2	3.8	864	13	CNS022LE
37	39.8	3.8	387	10	AW193322
38	39.8	3.8	427	10	BE048584
39	39.8	3.8	432	10	AW175624
40	39.8	3.8	441	10	AI818468
41	39.8	3.8	444	10	AW072844
42	39.8	3.8	450	10	AW273147
43	39.8	3.8	453	10	AW337946
44	39.8	3.8	492	13	AQ686347
45	39.8	3.8	505	10	AI952164

ALIGNMENTS

RESULT 1

BI411128
LOCUS 602962472F1 NCI_CGAP_Lu33 Mus musculus cdna clone IMAGE:5118185 5', mRNA sequence.
DEFINITION BI411128
ACCESSION BI411128.1 GI:15172051
VERSION EST.
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 973)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM11287 row: 0 column: 18
High quality sequence start: 32
High quality sequence stop: 912.
Location/Qualifiers
1..973
/organism="Mus musculus"
/strain="CZECH 11"
/db_xref="taxon:10090"
/clone="IMAGE:5118185"

```

/clone_lib="NCI_CGAP_Lu33"
/tissue_type="pooled lung tumors"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: p7T3D-Pac (Pharmacia) with a
modified polylinker; Site.1: NotI; Site.2: EcoRI; 1st
strand cDNA was prepared from mRNA obtained from pooled
lung tumors with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCGCTCTGTTTTTTTTTTTTT 3'] .
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified p7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Ronaldo. "
192 a 295 c 254 q 232 t
BASE COUNT

```

BASE COUNT	192 a	295 c	232 t	by Bento Soares and M. Fatima Bonaldo.	0
ORIGIN	constructed	295 c	232 t		
Query Match	30.6%	Score 320.4	DB 11	Length 973	
Best Local Similarity	77.1%	Pred. No. 1.8e-62			
Matches 390	Conservative	0	Mismatches 116	Indels	0
Gaps					0
QY	542	CCTCGAGTCTACCCAGAGCTTCCTTTTAAAGCTTAGAGCAAGTGAGAAAGATCCAGG	601		
DB	91	CCCTGCCTCTGCCCGAAGCTTCCTGCTTAAGTCCCTGGAGCAAGTGAGAAAGATCCAGG	150		
QY	602	CGGATGGGCGACGCGTCCAGGAAAGCTGTGCCACCTACAGCTGTCGCCACCCGAGG	661		
DB	151	CCAGCGGCTCGGTGCTGTGGAGCAGTTGTGCCACCTACAGCTGTGTACCCCGAGG	210		
QY	662	AGCTGGTCTGCTGGACACTCTCTGGGATCCCTTGGGCTCCCTGAGCAGCTGCCGCCA	721		
DB	211	AGCTGGTCTGCTGGGCACTCTCTGGGATCCCGAAGCTTCCCTGAGTGGCTGCTCTA	270		
QY	722	GCCAGGCCCTGACGTGGCAGGTGCTTGAGGCCAATCCATAGCGGCCTTTTCTCTACC	781		
DB	271	GCCAGGCCCTGACGACAGACAGATGGCTTAAGCAGCTCCACAGTGGCTTCGCTCTACC	330		
QY	782	AGGGCTCTCTGAGGCCCTGGAAGGATCTCCCGCAGTTGGTCCCACTTGGACAC	841		
DB	331	AAGTCTCTCGAGGCTCTATCGGGATTTTCCCTCGCCCTGGCCCCCACCCTTGGACTGC	390		
QY	842	TGCAGCTGGAGCTGCGGACTTTGCCACCAACATCTGCGACAGATGGAAGAACTGGAA	901		
DB	391	TTGAGCTGGATGTGCCAACTTTGCCACCACCACATCTGCGACGATGGAACCTAGGGG	450		
QY	902	TGCCCCCTGCCCCTGCAGCCCCACCAGGTTGCCATCCGCGCTTCGCTCTCTTCACG	961		
DB	451	TGCCCCCTTACTGTGCAGCCCCACACAGAGCGCATGCCAGCTTCACITCTGCTTCAGC	510		
QY	962	GCCGGGACAGAGGGTCTTAGTTGCTTCCATCTGCAGAGCTTCTGGAGGTGCTGAC	1021		
DB	511	GCCGGGACAGAGTGCTTGGCCATTTCTGCTACGTCAGGGCTTCTGGAGACGCTCGCC	570		
QY	1022	CGCTTCTACGCCACCTTGCCCGAGCC	1047		
DB	571	TTGCTCTGACCACTTGGCTAGACC	596		

RESULT	2	
LOCUS	CG548320	
DEFINITION	602575289f1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4703159 5', mRNA sequence.	
ACCESSION	CG548320	598 bp mRNA EST O4-APR-2001
VERSION	CG548320.1	GI:13546985
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS	1 (bases 1 to 598)	
TITLE	NIH-MGC http://mgc.nci.nih.gov/ . National Institutes of Health, Mammalian Gene Collection (MGC)	

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCMI541 row: b column: 24
High quality sequence stop: 597.

FEATURES	SOURCE
1. <i>General</i>	
2. <i>Specific</i>	
3. <i>Other</i>	
4. <i>Other</i>	
5. <i>Other</i>	
6. <i>Other</i>	
7. <i>Other</i>	
8. <i>Other</i>	
9. <i>Other</i>	
10. <i>Other</i>	
11. <i>Other</i>	
12. <i>Other</i>	
13. <i>Other</i>	
14. <i>Other</i>	
15. <i>Other</i>	
16. <i>Other</i>	
17. <i>Other</i>	
18. <i>Other</i>	
19. <i>Other</i>	
20. <i>Other</i>	
21. <i>Other</i>	
22. <i>Other</i>	
23. <i>Other</i>	
24. <i>Other</i>	
25. <i>Other</i>	
26. <i>Other</i>	
27. <i>Other</i>	
28. <i>Other</i>	
29. <i>Other</i>	
30. <i>Other</i>	
31. <i>Other</i>	
32. <i>Other</i>	
33. <i>Other</i>	
34. <i>Other</i>	
35. <i>Other</i>	
36. <i>Other</i>	
37. <i>Other</i>	
38. <i>Other</i>	
39. <i>Other</i>	
40. <i>Other</i>	
41. <i>Other</i>	
42. <i>Other</i>	
43. <i>Other</i>	
44. <i>Other</i>	
45. <i>Other</i>	
46. <i>Other</i>	
47. <i>Other</i>	
48. <i>Other</i>	
49. <i>Other</i>	
50. <i>Other</i>	
51. <i>Other</i>	
52. <i>Other</i>	
53. <i>Other</i>	
54. <i>Other</i>	
55. <i>Other</i>	
56. <i>Other</i>	
57. <i>Other</i>	
58. <i>Other</i>	
59. <i>Other</i>	
60. <i>Other</i>	
61. <i>Other</i>	
62. <i>Other</i>	
63. <i>Other</i>	
64. <i>Other</i>	
65. <i>Other</i>	
66. <i>Other</i>	
67. <i>Other</i>	
68. <i>Other</i>	
69. <i>Other</i>	
70. <i>Other</i>	
71. <i>Other</i>	
72. <i>Other</i>	
73. <i>Other</i>	
74. <i>Other</i>	
75. <i>Other</i>	
76. <i>Other</i>	
77. <i>Other</i>	
78. <i>Other</i>	
79. <i>Other</i>	
80. <i>Other</i>	
81. <i>Other</i>	
82. <i>Other</i>	
83. <i>Other</i>	
84. <i>Other</i>	
85. <i>Other</i>	
86. <i>Other</i>	
87. <i>Other</i>	
88. <i>Other</i>	
89. <i>Other</i>	
90. <i>Other</i>	
91. <i>Other</i>	
92. <i>Other</i>	
93. <i>Other</i>	
94. <i>Other</i>	
95. <i>Other</i>	
96. <i>Other</i>	
97. <i>Other</i>	
98. <i>Other</i>	
99. <i>Other</i>	
100. <i>Other</i>	

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4703159"
/clone_lib="NIH_MGC_77"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggccgcctgcgc); Site_2: SfiI (ggccattatgcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGCCATTATGCCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGGAGCGGCCGACATC-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.9
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."
113 a 202 c 163 q 120 t

```

BASE COUNT
ORIGIN

NOTE: THIS IS A NIMMGC LIBRARY.
 Cat. 202 c 163 g 120 t
 113 a
 29.9% Score 312.8; DB 11; Length 598;
 Similarity 99.4%;
 Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Query Match	29.9%	Score 312.8;	DB 11;	Length 598;
Best Local Similarity	99.4%;	Pred. No. 8.3e-61;		
Matches 314;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
QY 732	GCACCTGGCAGGCTGCTTGAGCCAACTCCATAGCGGCTTTTCTCTACCAAGGGGCTCCT	791		
Db 185	GAAGCTGGCAGGCTGCTTGAGCCAACTCCATAGCGGCTTTTCTCTACCAAGGGGCTCCT	244		
QY 792	GCAGGCGCTTGAAGGGGATCTCCCCGAGTTGGGTCCACCTTGGACACACTGCAGCTGGA	851		
Db 245	GCAGCGCTTGAAGGGGATCTCCCCGAGTTGGGTCCACCTTGGACACACTGCAGCTGGA	304		
QY 852	CGTCCGCACTTTGCCACCACTATCTGGCAGCAGATGGAGAAGCTGGGAATGCCCTGC	911		
Db 305	CGTCCGCACTTTGCCACCACTATCTGGCAGCAGATGGAGAAGCTGGGAATGCCCTGC	364		
QY 912	CTGTCAGCCCAACCCAGGGTGCCATGCGGCTCTGCCTTCAGCGCGCGGCAGG	971		
Db 365	CTGTCAGCCCAACCCAGGGTGCCATGCGGCTCTGCCTTCAGCGCGCGGCAGG	424		
QY 972	AGGGGTCCTAGTTGGCTTCCCATCTGCAGAGCTTCCTGAGGTTCTCTACCGCGTTCTACG	1031		
Db 425	AGGGGTCCTAGTTGGCTTCCCATCTGCAGAGCTTCCTGAGGTTCTCTACCGCGTTCTACG	484		
QY 1032	CCACCTTGCCCAAGCCC	1047		
Db 485	CCACCTTGCCCAAGCCC	500		

RESULT 3

RESULT	3	
BI328788		
LOCUS		
DEFINITION		30-JUL-2001
	BI328788	EST
	602984776Fl NCI_CGAP_L19	Mus musculus cDNA clone IMAGE:5137578 5',
	mRNA sequence.	
ACCESSION	BI328788	
VERSION	BI328788.1	GI:15013445
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	

REFERENCE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1 (bases 1 to 900)
TITLE	NIH-MGC http://mgc.nci.nih.gov/ .
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Jeffrey E. Green, M.D. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM1138 row: 9 column: 19 High quality sequence stop: 808.

```

FEATURES
source
1. .900
Location/Qualifiers
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_image="5137578"
/clone_lib="NCI_CGAP_Li9"
/lab_host="DH10B (T1 phage-resistant)"
/notes="Organ: liver; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.9 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
194 a 290 c 223 g 192 t 1 others
BASE COUNT
ORIGIN

```

Query Match 28.8%; Score 301.6; DB 11; Length 900;
Best Local Similarity 85.1%; Pred. NO. 3.1e-58;
Matches 337; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY	1	ATGGAGCTGACTGAATTTGCTCTCTGCTGGTCACTGCTTCTCCTTAAGTCCCAAGGCTTAAGCGTG	60
Db	191	ATGGAGCTGACTGATTTGCTCTCTGCGGGCAGTCATCTTCTTTCGATGTGGCAAGACTAACTCTG	250
QY	61	TCCAGCCCGGCTCCTCTGCTGTTGTGACCTCCAGATCCCTCAGTAAACTGCTTCGTGACTCC	120
Db	251	TCCAGCCCGGTAGCTCCTGCCTGTGACCCGAGACTCCTAATAAAGTGTGCGTGACTCC	310
QY	121	CATGTCCTTTACACGAGACTGAGCCAGTGCCTCCAGAGGTTTACCCCTTTGCGCTACACCTGTC	180
Db	311	CACCTCCTTCACAGCGCACTGAGTCAGTGTCCGACGTCGACCCTTTGCTATCCCTGTT	370
QY	181	CTGCTGCCTGCTGTGACCTTTAGCTTGGAGAAATGGAACCCAGATGAGAGAGACCCAG	240
Db	371	CTGCTGCCCTGCTGTGGACTTTAGCCTTGGAGAAATGGAACCCAGACGGAACAGAGACAAG	430
QY	241	GCACAGACATTCCTGGGAGCAGTGACCCCTTCTGCTGGAGGGAGTGATGCACACAGGGGA	300
Db	431	GCACAGACATTCCTAGGGGAGTGTCCTTCTACTGGAGGGAGTGATGCACACAGGGA	490
QY	301	CAACTGGGACCCACTTTCCTCTCATCCCTCCTGGGGCAGCTTTCTTGGACAGGTCCGCTCTC	360
Db	491	CAGTTGGAAACCCCTCCCTGCCTCTCATCCCTCCTGGGACAGCTTCTTGGGACAGTTTCGCCCTC	550
QY	361	CTCCTTGGGGCCCTGCAGAGCCCTCCTTTGGAACCCAG	396
Db	551	CTCTTGGGGCCCTGCAGGCCCTCCTTAGGAACCCAG	586

RESULT	4	
BE485194		
LOCUS	BE485194	507 bp mRNA
DEFINITION	172130 BARC 5BOV Bos taurus CDNA 5', mRNA sequence.	EST 28-AUG-2000
ACCESSION	BE485194	
VERSION	BE485194.1	GI:9604727
KEYWORDS	EST.	
SOURCE	COW.	

ORGANISM	Bos taurus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
AUTHORS	1 (bases 1 to 507) Sonstegard,T.S., Capuco,A.V., Van Tassel,C.P., Ashwell,M.S. and Wells,K.D.
TITLE	Mapping of Expressed Sequence Tags from a normalized bovine mammary gland cDNA library
JOURNAL	Unpublished (2000)
COMMENT	Contact: Sonstegard TS USDA, ARS, Beltsville Agricultural Research Center Bldg. 200 Rm 2A, Beltsville, MD 20705, USA Tel: 301 504 8416 Fax: 301 504 8414 Email: tats@psi.barc.usda.gov Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -mismatch 12 options.

FCN primers
 FORWARD: AGGAACACAGCTATGACCAT
 BACKWARD: GTTTCGCCAGTCAACGACG
 Plate: 134 row: J column: 17
 Seq primer: ATTTAGGTGACACTATAG.

```

FEATURES
source
Location/Qualifiers
1. .507
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="BARC 580V"
/tissue_type="pooled"
/lab_host="DH10B"
/notes="Vector: pCMV SPORT6; site_1: XbaI; Site_2: XhoI;
Library made from pooled mRNA isolated from mammary
tissues at eight physiological, developmental, and disease
states."

```

		States:	
BASE COUNT	97 a	170 c	82 t
ORIGIN			
		158 g	

Query Match	28.7%	Score 300.6;	DB 10;	Length 507;
Best Local Similarity	84.1%	Pred. No. 4.6e-58;		
Matches 339; Conservative	0;	Mismatches 64;	Indels 0;	Gaps 0;

630	GTGTGCCACCTACAAGTGTGGCCACCCCGAGGAGCTGTGTCTCTCGGACACTCTCTGGG	689
104	GTGTGGCCGCCACAAGCTGTGGCCACCCCGAGGAGCTGTGTCTCTCGGACACTCTCTGGG	163
690	CATCCCTGGGCTCCCTGAGCAGCTGCCCCAGCCAGGCCCTCGACGTGSCAGGCTGCTT	749
164	CATCCCCCAGGCTCCCTTAAGCAGCTGTCTCCAGCCAGTCCCTCGACCTGACGAGCTGCCT	223
750	GAGCCAACTCCATAGCGGGCTTTTCTCTACCAGGGGCTCTCTGCAGGCCCTTGGAAAGGAT	809
224	GAACCAACTACAGGGCGGCTTCTTCTCTACAGGGGCTCTCTGCAGGCCCTGCGGGGCAT	283
810	CTCCCCCGAGTTGGGTCCACCTTGGACACACTGACAGCTGGAGCTGCGCGACTTTGCCAC	869
284	CTCCCCAGAGCTGGCCCCACCTTGGACACACTGACAGCTGGAGCTGACTTTGCCAC	343
870	CACCATCTGGCAGCAGATGGAGAACTGGGAATGCCCTGCCCTCGACGCCACCCAGG	929
344	GAACATCTGGCTGCAGATGAGGACCTGGGGGGGGCCCCCGCTGTGCAGCCACCCAGGG	403
930	TGCATCGCGGCTTTCGCTCTCTTTTCAGCGCCGGCAGGAGGGGTCTTACTTTGCCCTC	989
404	CGCATCGCCGACCTTCACTTCAGCTTCCAAACCCAGCAGGAGGGGTCTGTTGCTTC	463
990	CCATCTGCAGAGCTTCTCGAGGTGCTGTAACCGGTTCTACGC	1032
464	CCAGCTGCATCGTTTCTTGGAGCTGGCATAACCTGGCCCTGGCG	506

RESULT 5

D	b	358	CACCTCCTTCACAGCCGACTTGAGTCA	GTTGCCGACGTCCCGACGTCGACCCCTTTGTCATCCCTGTT	417					
Q	y	181	CTGCTGCCTGCTGGACACTTTAGCTT	GGGAAGTAATGGAAACCAGATGGAGGACCAAG	240					
D	b	418	CTGCTGCCTGCTGGACACTTTAGCTT	GGGAAGTAATGGAAACCAGACGACGACGAAG	477					
Q	y	241	GCACAGGACATTCTGGGACAGTAGTAC	CCCTTCTGCTGGAGGAGTGATGGCAGCAGGGGA	300					
D	b	478	GCACAGGACATTCTAGGGCAGTAGTCC	CTTCTACTGGAGGAGTGATGGCAGCACGAGA	537					
Q	y	301	CAACTGGGACCCACTTGGCTCTCAT	TCCCTCTCTGGGCGAGCTTCTGGACAGAGTCCGTC	360					
D	b	538	CAGTTGGAAACCTCTGCTCTCAT	TCCCTCTGCGGACAGCTNTCTGGGCAAGTTCCGCTC	597					
Q	y	361	CTCCTTTGGGCCCCTGCAGAGCCCT	CTTGGAAACCCAG	396					
D	b	598	CTCTTGGGGGCCCTGCAGGGCCCT	CCTAGGAACCCAG	633					
RESULT 6										
AA538257										
LOCUS	AA538257		609 bp	mRNA	EST 29-JUL-1997					
DEFINITION	vJ03407.r1 Barstead mouse pooled organs MPLR84 Mus musculus cDNA clone IMAGE:920628 5' similar to gb:L34169 Mus musculus thromboopoietin mRNA, complete cds (MOUSE);, mRNA sequence.									
ACCESSION	AA538257									
VERSION	AA538257.1 GI:2284250									
KEYWORDS	EST.									
SOURCE	house mouse.									
ORGANISM	Mus musculus									
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.									
AUTHORS	1 (bases 1 to 609)									
TITLE	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Willson,R. and Waterston,R.									
	The WashU-HMI Mouse EST Project									
	Unpublished (1996)									
	Contact: Marra M/Mouse EST Project									
	WashU-HMI Mouse EST Project									
COMMENT	Washington University School of Medicinep									
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108									
	Tel: 314 286 1800									
	Fax: 314 286 1810									
	Email: mouseest@watson.wustl.edu									
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.										
MGI:532844										
Seq primer: -28ml3 rev2 ET from Amersham										
High quality sequence stop: 492.										
FEATURES	Location/Qualifiers									
	1..609									
	/organism="Mus musculus"									
	/strain="FVB/N"									
	/db_xref="taxon:10090"									
	/clone="IMAGE:920628"									
	/clone_lib="Barstead mouse pooled organs MPLRB4"									
	/sex="mixed"									
	/tissue_type="pooled organs"									
	/dev_stage="7 day"									
source	/lab_host="DH10B"									
	/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI ; Site_2: NotI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'- TGTAGCAATCTGAAGTGGAGCGCGCTTTTTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors [GTTGATTCGGTACC], digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library constructed by Bob Barstead."									
	126	a	204	c	157 g 122 t					
BASE COUNT										
ORIGIN										

PCR Primers

FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCCTCAGCTCAGCAG
 Plate: 54 row: J column: 23
 Seq primer: ATTAGGTGACACTATAG.

FEATURES

source

Location/Qualifiers
 1..464
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /clone_lib="MARC 2BOV"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
 Library made from pooled tissue from testis, thymus,
 semitendinosus muscle, longissimus muscle, pancreas,
 adrenal, and endometrium."

BASE COUNT
 ORIGIN

97 a 152 c 128 g 87 t

Query Match 22.0%; Score 230; DB 10; Length 464;
 Best Local Similarity 85.5%; Pred. No. 4.3e-42;
 Matches 284; Conservative 0; Mismatches 35; Indels 13; Gaps 2;
 QY 203 GCTTGGGAGATGGAACCCAGATGAGGAGACCAAGGACAGACATTCCTGGAGCAG 262
 Db 1 GCTTGGGAGATGGAACCCAGATGAGGAGACCAAGGACAGACATTCCTGGAGCAG 262
 QY 263 TGACCCCTTCCTGCTGGAGGAGTGGGAGCAGCGGGGACAACTGGGACCCACTTGCCTCT 322
 Db 61 CGACCCCTTCCTGCTGGAGGAGTGGGAGCAGCGGGGACAACTGGGACCCACTTGCCTCT 322
 QY 323 CATCCCTTCCTGGGAGCTTCCTGGACAGTCCGCTCCTCTCTGGGGCCCTGCAGAGC 382
 Db 121 CATCCCTTCCTGGGAGCTTCCTGGGAGTCCGCTCCTCTCTGGGGCCCTGCAGAGC 382
 QY 383 TCCTTGGACCCAGCTTCTCCACAGGCGAGGACACAGCTCACAAGGATCCCAATGCCA 442
 Db 181 TCCTTGGACCCAGCTTCTCCACAGGCGAGGACACAGCTCACAAGGATCCCAATGCCA 442
 QY 443 TCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGTGGCTTCCCTGATGCTGT-AGGA 501
 Db 229 TCTTCTGAGCTTCCACAGCTGCTCCGAGGAAAGTGGCTTCCCTGATGCTGT-AGGA 501
 QY 502 GGGTCCACCTCTGCTGACGCGGGGCCCAAC 533
 Db 289 GGGCCACCTCTGTGCCAAGCGGGGCCCAAC 320

RESULT 9

BF848766

LOCUS

BF848766 327 bp mRNA

QV0-EN0102-081100-458-g06 EN0102 Homo sapiens EST 16-JAN-2001

ACCESSION BF848766 QV0-EN0102-081100-458-g06 EN0102 Homo sapiens cdNA, mRNA sequence.

VERSION BF848766.1 GI:122335903

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Dias Neto,B., Garcia Correa,R., Verjovsky-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.

TITLE

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

JOURNAL

MEDLINE

COMMENT

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl-QV0&t2-QV0-EN0102-
 081100-458-g06&t3-2000-11-08&t4-1)

Seq primer: puc 18 forward
 High quality sequence stop: 327.

Location/Qualifiers
 1..327

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="EN0102"
 /dev_stage="Adult"

/note="Organ: lung, normal; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

BASE COUNT 61 a 87 c 103 g 76 t

ORIGIN

Query Match 13.0%; Score 135.8; DB 11; Length 327;
 Best Local Similarity 89.6%; Pred. No. 8.4e-21;
 Matches 146; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 726 GGCCTCGAGCTGGCAGGCTGCTTGAGCCAACTCCATAGCGCCTTTCTCTACCAAGG 785
 Db 139 GGCCTCGAGCTGGCAGGCTGCTTGAGCCAACTCCATAGCGCCTTTCTCTACCAAGG 785

QY 786 GCTCTCGAGCCCTGGAAGGATCTCCCGAGTTGGTCCCACTTGGACACACTGCA 845
 Db 199 GCTCTCGAGCCCTGGAAGGATCTCCCGAGTTGGTCCCACTTGGACACACTGCA 258

QY 846 GCTGAGCTGCGGACTTTGCCACCACTTGGCAGCAGATG 888
 Db 259 GCTGAGCTGCGGACTTTGCCACCACTTGGCAGCAGATG 301

RESULT 10
 AZ025993 449 bp DNA GSS 25-FEB-2000

LOCUS RPCI-23-343116.TJ RPCI-23 Mus musculus genomic clone RPCI-23-343116

DEFINITION , DNA sequence.

ACCESSION AZ025993

VERSION AZ025993.1 GI:7101377

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE 1 (bases 1 to 449)

AUTHORS Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
 B., Levins,M., Mcann,S., Isegaye,G., Geer,K., Krol,M., de Jong,P.
 and Fraser,C.M.

Mouse BAC End Sequences from Library RPCI-23

Unpublished (1999)

Other_GSSs: RPCI-23-343116.TV

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC

library availability, please contact Pieter de Jong (pieter@jeong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>) or from Resea ch Genetics (info@resgen.com). BAC end page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
 Plate: 343 row: I column: 16
 Seq primer: SP6
 Class: BAC ends.

FEATURES

source
 Location/Qualifiers
 1..449
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="RPCI-23-343I16"
 /clone_lib="RPCI-23"
 /sex="Female"
 /lab_host="DH10B"
 /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1: ECORI; Site_2: ECORI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methyase. Size selected DNA was cloned into the pBACe3.6 vector at the ECORI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 90 a 121 c 108 g 129 t 1 others

ORIGIN

Query Match 12.9%; Score 135.2; DB 13; Length 449;
 Best Local Similarity 84.4%; Pred. No. 1.2e-20;
 Matches 152; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
 QY 228 GGAGGAGCAGGAGCAGGAGCAGTCTGGGAGCAGTGAACCTCTGCTGGAGGAGTGCAT 287
 Db 126 GGAGGAGCAGGAGCAGGAGCAGTCTGGGAGCAGTGAACCTCTGCTGGAGGAGTGCAT 185
 QY 288 GGCAGCAGGGGAGCACTGGGAGCCACTTGCCTCTCATCCCTCTGGGAGCAGTCTTCTGG 347
 Db 186 GGCAGCAGGAGCAGTGGAAACCTCTGCTCTCATCCCTCTGGGAGCAGTCTTCTGG 245
 QY 348 ACAGGTGGTCTCTCTCTGGGGCCCTGGCAGAGCCTCTTGGAAACCCAGCTTCCTCCACA 407
 Db 246 GCAGGTGGCTCTCTTGGGGCCCTGGCAGGGCTCTAGGAACCCAGGTAAGTCCCCA 305

RESULT 11

AZ352219/c 419 bp DNA GSS 29-SEP-2000
 LOCUS 1M0090P05R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 DEFINITION clone UUGC1M0090P05 R, DNA sequence.
 ACCESSION AZ352219
 VERSION AZ352219.1 GI:10431456
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 419)
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00
 Plate: 0090 row: P column: 05
 Seq primer: CACACAGGAACACCATGACG
 Class: plasmid ends
 High quality sequence stop: 419.
 Location/Qualifiers

FEATURES

source
 Location/Qualifiers
 1..419
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0090P05"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Jax DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gil4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 100 a 98 c 124 g 97 t
 ORIGIN

Query Match 10.4%; Score 109.4; DB 13; Length 419;
 Best Local Similarity 76.6%; Pred. No. 8.4e-15;
 Matches 134; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
 QY 360 CCTCTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCTCCACAGGCGAGGACCAC 419
 Db 292 CCTACTCACTGGCTCAGGCGCCATCCTCTGCCCTCAGCTTCTCTACAGGCGAGGACCAC 233
 QY 420 AGCTCAAGAATCCCAATGCCATCTTCTCAGCTTCCAAACCTGCTCCGAGGAAAGGT 479
 Db 232 AGCTCAAGAAGCCCAATGCCCTCTTCTTTCAGCTTCCAAACCTGCTCCGAGGAAAGGT 173
 QY 480 GCGTTTCTGATCTTGTAGGAGGTCACCCCTCTGCTAGCGGGCGGCGCAACA 534
 Db 172 GCGCTTCTGCTTGTGTAGAGTCCACCCCTCTGCTGTCAGAGGAGCCCTGCCA 118

RESULT 12

AZ792988 708 bp DNA GSS 16-FEB-2001
 LOCUS 2M0046I05F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 DEFINITION clone UUGC2M0046I05 F, DNA sequence.
 ACCESSION AZ792988
 VERSION AZ792988.1 GI:12937676
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 708)
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL
COMMENT

Unpublished (2000)
Contact: Robert B. Weiss
University of Utah
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0046 row: 1 column: 05
Seq primer: CGTGTAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 708.
Location/Qualifiers
1. .708
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0046105"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male); was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (G1/4732114/gb/AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

FEATURES

source

1. .708
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0046105"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male); was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (G1/4732114/gb/AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
ORIGIN

156 a 213 c 147 g 192 t

Query Match

Best Local Similarity 8.3%; Score 86.4; DB 13; Length 708;
Matches 111; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy 16 TTGCTCCTCGTGGTGCATGCTTCTCCCTAACGCAAGGCTAACGGTGTCCAGCCGGCTCCT 75

Db 1 TTGCTCCTCGCGCCATGCTTCTTGCAAGTGGCAAGACTAACTCTGTCCAGCCCGTAGCT 60

Qy 76 CTTGCTTTGACCTCGGACCTCCTCAGTAAACTGCTTTCGTGACTCCCATGCTTTCACAGC 135

Db 61 CTTGCTTTGACCTCGGACCTCCTCAGTAAACTGCTTTCGTGACTCCCATGCTTTCACAGC 120

Qy 136 AGACTGAGCCAGTGCACAGAGTTTACCCCTTT 167

Db 121 CGACTGCTGAGCAACCCCAAGCCCTTGTCTT 152

RESULT 13

A0888740

LOCUS A0888740 752 bp DNA GSS 10-NOV-1999
DEFINITION HS_3106_B1_B04_MR CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone Plate=3106 Col=7 Row=D, DNA sequence.

ACCESSION

A0888740

VERSION A0888740.1 GI:6344930

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE

AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 752)
Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D., and
Hood, L.

TITLE

JOURNAL

MEDLINE

COMMENT

Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)

Contact: Mahairas GG, Wallace JC, Hood L

High Throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Clones may be purchased from Research Genetics (info@resgen.com).

BAC end Web Server: http://www.htsc.washington.edu

Plate: 3106 row: D column: 7

Seq primer: M13 Reverse

Class: BAC ends

High quality sequence stop: 752.

FEATURES

source

1. .752

Location/Qualifiers

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="Plate=3106 Col=7 Row=D"

/clone_lib="CIT Approved Human Genomic Sperm Library D"

/sex="male"

/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in

E-Coli DH10B"

BASE COUNT 181 a 219 c 177 g 175 t

ORIGIN

Query Match

Best Local Similarity 7.4%; Score 77.8; DB 13; Length 752;

Matches 123; Conservative 0; Mismatches 57; Indels 1; Gaps 1;

Qy 13 GAATTCCTCCTCGTGGTGCATGCTTCTCCCTAACGCAAGGCTAACGGTGTCCAGCCCGGCT 72

Db 509 GAATTCCTCCTCGTGGTGCATGCTTCTCCCTAACGCAAGGCTAACGGTGTCCAGCCCGGCT 568

Qy 73 CCTCCTCCTTGTGACCTCCGAGTCCCTCAGTAAACTGCTCGTACCTCCCATGCTCTTCAAC 132

Db 569 CCTACTGTTGTGAACCTCCATCTCCTCAG-GAAGTGTTCGGGACTCCCATGGCTCACA 627

Qy 133 AGCAGACTGAGCCAGTGCACAGAGTTTCAACCTTTGCCTACACCTGTCCTGCTGCTGCT 192

Db 628 GCAAACTGGTGAGATCTCTCAACAGTATCCTCTTATCCGCTTAACCTGTGAAGACCACTACT 687

Qy 193 G 193

Db 688 G 688

RESULT 14

BE377769

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

BE377769 589 bp mRNA EST 21-JUL-2000
60122920F1 NC1_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3593885 5',
mRNA sequence.

BE377769

BE377769.1 GI:9323134

EST.

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 589)

NIH-MGC http://mgs.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Fri Apr 5 13:03:31 2002

us-09-680-514-4.rst

Page 10

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2002, 10:29:29 ; Search time 2520.15 Seconds
(without alignments)
7089.437 Million cell updates/sec

Title: US-09-680-514-6
Perfect score: 1083
Sequence: 1 ATGGAGCTGACTGAATTGCT.....TACGCCACCTTGCCAGCCC 1083

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 1472140 seqs, 8248589755 residues
Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl : *
1: gb_ba : *
2: gb_htg : *
3: gb_in : *
4: gb_om : *
5: gb_ov : *
6: gb_pat : *
7: gb_ph : *
8: gb_pl : *
9: gb_pr : *
10: gb_ro : *
11: gb_sts : *
12: gb_sy : *
13: gb_un : *
14: gb_vi : *
15: em_ba : *
16: em_fun : *
17: em_hum : *
18: em_in : *
19: em_om : *
20: em_or : *
21: em_ov : *
22: em_pat : *
23: em_ph : *
24: em_pl : *
25: em_ro : *
26: em_sts : *
27: em_sy : *
28: em_un : *
29: em_vi : *
30: em_htgo_hum : *
31: em_htgo_inv : *
32: em_htgo_rod : *
33: em_htg_hum : *
34: em_htg_inv : *
35: em_htg_rod : *
36: em_htg_other : *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	523.2	48.3	555	22	E11965	E11965 Human cDNA
2	523.2	48.3	1059	6	AR103959	AR103959 Sequence
3	523.2	48.3	1062	6	AR087133	AR087133 Sequence
4	523.2	48.3	1062	6	AR087571	AR087571 Sequence
5	523.2	48.3	1062	6	I49760	I49760 Sequence 3
6	523.2	48.3	1062	9	HUMTHROMB	L36052 Human throm
7	523.2	48.3	1086	6	E16669	E16669 cDNA encodi
8	523.2	48.3	1267	6	E12182	E12182 Human cDNA
9	523.2	48.3	1267	6	E12214	E12214 Human cDNA
10	523.2	48.3	1342	6	AR008878	AR008878 Sequence
11	523.2	48.3	1342	6	AR023468	AR023468 Sequence
12	523.2	48.3	1342	6	AR087573	AR087573 Sequence
13	523.2	48.3	1342	6	I85129	I85129 Sequence 1
14	523.2	48.3	1691	9	D32047	D32047 Homo sapien
15	523.2	48.3	1721	6	E16668	E16668 cDNA encodi
16	523.2	48.3	1721	22	E11961	E11961 Human cDNA
17	523.2	48.3	1795	6	I33525	I33525 Sequence 1
18	523.2	48.3	1795	9	HUMMLCML	L33410 Human c-mpl
19	522	48.2	525	6	E02574	E02574 DNA encodin
20	521.6	48.2	605	6	AR008879	AR008879 Sequence
21	521.6	48.2	605	6	AR087574	AR087574 Sequence
22	521.6	48.2	605	6	I85130	I85130 Sequence 3
23	520	48.0	1341	9	HSU11025	U11025 Human megak
24	518.4	47.9	1062	9	HSU59493	U59493 Human throm
25	518.4	47.9	1062	9	HSU59494	U59494 Human throm
26	515.6	47.6	525	6	E02575	E02575 DNA encodin
27	514	47.5	525	6	E02576	E02576 DNA encodin
28	489.2	45.2	1050	22	E11559	E11559 Human cDNA
29	486.8	44.9	525	6	AR024358	AR024358 Sequence
30	486.8	44.9	525	6	AR091731	AR091731 Sequence
31	486.8	44.9	525	6	E01731	E01731 cDNA encodi
32	486.8	44.9	525	6	E02573	E02573 DNA encodin
33	486.8	44.9	525	6	E07164	E07164 cDNA encodi
34	486.8	44.9	525	6	E15131	E15131 Human mRNA
35	486.8	44.9	525	6	I71150	I71150 Sequence 1
36	486.8	44.9	525	6	I83709	I83709 Sequence 1
37	486.8	44.9	525	22	E09431	E09431 cDNA encodi
38	486.8	44.9	660	9	HUMGCSF	M17706 Human granu
39	486.8	44.9	1498	9	HSGCSFR1	X03655 Human mRNA
40	486.8	44.9	1521	6	E01219	X03655 Human mRNA
41	486.8	44.9	1525	6	I05333	I05333 Sequence 11
42	486.8	44.9	1525	6	I08669	I08669 Sequence 1
43	486.8	44.9	1525	6	I09205	I09205 Sequence 5
44	485.2	44.8	522	6	E01631	E01631 DNA encodin
45	485.2	44.8	522	6	E08529	E08529 DNA encodin

ALIGNMENTS

RESULT 1
E11965
ID E11965 standard; RNA; HUM; 555 BP.
XX
AC E11965;
XX
SV E11965.1
XX
DT 08-OCT-1997 (Rel. 52, Created)
DT 02-SEP-2000 (Rel. 65, Last updated, Version 2)
XX
DE Human cDNA encoding thrombopoietin.
XX
KW JP 1996228781-A/6.
XX
OS Homo sapiens (human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX
RN [1]

1-555
 RA Miyazaki H., Kato H., Ogami K., Iwamatsu A., Akahori H., Kuroki R.,
 RA Shimizu T., Muto T.;
 RT "DNA CODING FOR PROTEIN HAVING TPO ACTIVITY";
 RL Patent number JP1996228781-A/6, 10-SEP-1996.
 RL KIRIN BREWERY CO LTD.
 XX
 CC OS Homo sapiens (human)
 CC PN JP 1996228781-A/6
 CC PD 10-SEP-1996
 CC PF 14-FEB-1995 JP 1995063298
 CC PR 14-FEB-1994 JP 94P 39090, 25-MAR-1994 JP 94P -79842,
 CC PR 01-JUN-1994 JP 94P 155126, 15-JUN-1994 JP 94P 167328,
 CC PR 17-AUG-1994 JP 94P 227159, 17-AUG-1994 JP 94P 193169,
 CC PR 01-NOV-1994 JP 94P 304167, 01-DEC-1994 JP 94P 298669,
 CC PR 28-DEC-1994 JP 94P 341200
 CC PI MIYAZAKI HIROSHI, KATO HISASHI, OGAMI KINYA, IWAMATSU AKIHIKO,
 CC PI AKAHORI HIRONORI, KUROKI RYOTA, SHIMIZU TOSHIYUKI,
 CC PI MUTO TAKANORI
 CC PC C12N15/09,A61K38/00,C07K14/52,C12N1/21,C12N5/10,C12P21/02,
 CC PC (C12N1/21,
 CC PC C12R1:19), (C12N5/10,C12R1:91), (C12P21/02,C12R1:19), (C12P21/02,
 CC PC C12R1:91);
 CC CC strandedness: Double;
 CC CC topology: Linear;
 CC FH Key Location/Qualifiers
 CC FH source 1..555
 CC FT /organism="Homo sapiens"
 CC FT /tissue_type="liver",
 CC FT CDS 1..555
 CC FT /product="thrombopoietin"
 CC FT sig_peptide 1..63
 XX
 FH Key Location/Qualifiers
 FH source 1..555
 FH /db_xref="taxon:9606"
 FH /organism="Homo sapiens"
 SQ Sequence 555 BP; 106 A; 183 C; 145 G; 121 T; 0 other;

Query Match 48.3%; Score 523.2; DB 22; Length 555;
 Best Local Similarity 99.4%; Pred. No. 5.9e-99;
 Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 ATGGAGCTGACTGAATTCCTCGTGGTCACTCTCTTAACCTGCAAGGCTAACGCTG 60
 DB 1 ATGGAGCTGACTGAATTCCTCGTGGTCACTCTCTTAACCTGCAAGGCTAACGCTG 60
 QY 61 TCCAGCCCGGCTCCTCTCTGCTGTGACCTCCGAGTCCAGTCCCTAGTAACTGCTTCGTAAGTCC 120
 DB 61 TCCAGCCCGGCTCCTCTCTGCTGTGACCTCCGAGTCCAGTCCCTAGTAACTGCTTCGTAAGTCC 120
 QY 121 CATGTCTCTACAGCAGACTGAGCCAGTCCGAGAGGTTCCACCTTTGCTTACACCTGTC 180
 DB 121 CATGTCTCTACAGCAGACTGAGCCAGTCCGAGAGGTTCCACCTTTGCTTACACCTGTC 180
 QY 181 CTGCTGCTCTGTGGACTTTAGCTTGGGAGATGGAAACCCAGATGGAGGAGACCAAG 240
 DB 181 CTGCTGCTCTGTGGACTTTAGCTTGGGAGATGGAAACCCAGATGGAGGAGACCAAG 240
 QY 241 GCACAGGACATTCCTGGGAGCAGTACCTTCTGCTGGAGGAGTATGGCAGCGGGA 300
 DB 241 GCACAGGACATTCCTGGGAGCAGTACCTTCTGCTGGAGGAGTATGGCAGCGGGA 300
 QY 301 CAACCTGGGACCCACTTGCCTCTATCCCTCTGCGGGCAGCTTCTTGGACAGGTCCTGTC 360
 DB 301 CAACCTGGGACCCACTTGCCTCTATCCCTCTGCGGGCAGCTTCTTGGACAGGTCCTGTC 360
 QY 361 CTCCTTGGGGCCCTGCAGAGCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCAACA 420
 DB 361 CTCCTTGGGGCCCTGCAGAGCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCAACA 420
 QY 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
 DB 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
 QY 481 CGTTTCTCTGATGCTGTAGGAGGTCACACCTCTGCTCGAGGGTGGC 528

Db 361 CTCCTTGGGGCCCTGCAGAGCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCAACA 420
 QY 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
 Db 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
 QY 481 CGTTTCTCTGATGCTGTAGGAGGTCACACCTCTGCGTCAAGGGTGGC 528
 Db 481 CGTTTCTCTGATGCTGTAGGAGGTCACACCTCTGCGTCAAGGGTGGC 528
 RESULT 2
 ARI03959
 LOCUS AR103959 1059 bp DNA PAT 14-FEB-2001
 DEFINITION Sequence 57 from patent US 6093395.
 ACCESSION AR103959
 VERSION AR103959.1 GI:12816667
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1059)
 AUTHORS Bauer, S. Christopher, Abrams, M. Allen, Braford-Goldberg, S. Ruth,
 Caparon, M. Helena, Easton, A. Michael, Klein, B. Kure, McEarn, J. P.,
 Orlins, P. O., Paik, K. and Thomas, J. W.
 TITLE Co-administration of interleukin-3 mutant polypeptides with CSF's
 JOURNAL for multi-lineage hematopoietic cell production
 FEATURES Patent: US 6093395-A 57 25-JUL-2000;
 source Location/Qualifiers
 1..1059
 BASE COUNT 221 a 375 c 236 g 227 t
 ORIGIN

Query Match 48.3%; Score 523.2; DB 6; Length 1059;
 Best Local Similarity 99.4%; Pred. No. 5e-99;
 Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 ATGGAGCTGACTGAATTCCTCGTGGTCACTCTCTTAACCTGCAAGGCTAACGCTG 60
 DB 1 ATGGAGCTGACTGAATTCCTCGTGGTCACTCTCTTAACCTGCAAGGCTAACGCTG 60
 QY 61 TCCAGCCCGGCTCCTCTCTGCTGTGACCTCCGAGTCCCTAGTAACTGCTTCGTAAGTCC 120
 DB 61 TCCAGCCCGGCTCCTCTCTGCTGTGACCTCCGAGTCCCTAGTAACTGCTTCGTAAGTCC 120
 QY 121 CATGTCTCTACAGCAGACTGAGCCAGTCCGAGAGGTTCCACCTTTGCTTACACCTGTC 180
 DB 121 CATGTCTCTACAGCAGACTGAGCCAGTCCGAGAGGTTCCACCTTTGCTTACACCTGTC 180
 QY 181 CTGCTGCTCTGTGGACTTTAGCTTGGGAGATGGAAACCCAGATGGAGGAGACCAAG 240
 DB 181 CTGCTGCTCTGTGGACTTTAGCTTGGGAGATGGAAACCCAGATGGAGGAGACCAAG 240
 QY 241 GCACAGGACATTCCTGGGAGCAGTACCTTCTGCTGGAGGAGTATGGCAGCGGGA 300
 DB 241 GCACAGGACATTCCTGGGAGCAGTACCTTCTGCTGGAGGAGTATGGCAGCGGGA 300
 QY 301 CAACCTGGGACCCACTTGCCTCTATCCCTCTGCGGGCAGCTTCTTGGACAGGTCCTGTC 360
 DB 301 CAACCTGGGACCCACTTGCCTCTATCCCTCTGCGGGCAGCTTCTTGGACAGGTCCTGTC 360
 QY 361 CTCCTTGGGGCCCTGCAGAGCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCAACA 420
 DB 361 CTCCTTGGGGCCCTGCAGAGCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCAACA 420
 QY 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
 DB 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
 QY 481 CGTTTCTCTGATGCTGTAGGAGGTCACACCTCTGCTCGAGGGTGGC 528

Db 481 CGTTTCTGATGCTGTAGGAGGTCCACCTCTGCGTCAGGCGGGCC 528

RESULT 3

LOCUS AR087133 1062 bp DNA PAT 07-SEP-2000
DEFINITION Sequence 3 from patent US 5986049.
ACCESSION AR087133
VERSION AR087133.1 GI:10013899
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1062)

AUTHORS Forstrom,J.W., Lofton-Day,C.E. and Lok,S.
TITLE Purified thrombopoietin and method of making it
JOURNAL Patent: US 5986049-A 3 16-NOV-1999;
FEATURES Location/Qualifiers
source 1..1062

BASE COUNT 223 a 375 c 236 g 228 t
ORIGIN

Query Match 48.3%; Score 523.2; DB 6; Length 1062;
Best Local Similarity 99.4%; Pred. No. 5e-99;
Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATTGCTCTCGTGTGCTATGCTTCTCTTAAGCTGAAGGCTAACGCTG 60
|||||
Db 1 ATGGAGCTGACTGAATTGCTCTCGTGTGCTATGCTTCTCTTAAGCTGAAGGCTAACGCTG 60
|||||
QY 61 TCCAGCCGGCTCTCGTGTGCTGAGCTCGGAGTCTCAGTAAGTCTGCTGACTCC 120
|||||
Db 61 TCCAGCCGGCTCTCGTGTGCTGAGCTCGGAGTCTCAGTAAGTCTGCTGACTCC 120
|||||
QY 121 CATGCTCTTACAGCAGACTGAGCCAGTGCAGAGGTTACCCCTTTGCCCTACACCTGTC 180
|||||
Db 121 CATGCTCTTACAGCAGACTGAGCCAGTGCAGAGGTTACCCCTTTGCCCTACACCTGTC 180
|||||
QY 181 CTGCTGCTGTGGAGCTTTAGCTTGGGAGATGGAACCCAGATGGAGGAGACCAAG 240
|||||
Db 181 CTGCTGCTGTGGAGCTTTAGCTTGGGAGATGGAACCCAGATGGAGGAGACCAAG 240
|||||
QY 241 GCACAGACATCTGGGAGCAGTCAACCTTCTGCTGGAGGAGTGATGGCAGCAGCGGA 300
|||||
Db 241 GCACAGACATCTGGGAGCAGTCAACCTTCTGCTGGAGGAGTGATGGCAGCAGCGGA 300
|||||
QY 301 CAACTGGGACCCACTTGCCTCTCATCCCTCTGCTGGGAGGCTTTCTGGACAGGTCGCTC 360
|||||
Db 301 CAACTGGGACCCACTTGCCTCTCATCCCTCTGCTGGGAGGCTTTCTGGACAGGTCGCTC 360
|||||
QY 361 CTCCTTGGGCGCTGCAGAGCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCA 420
|||||
Db 361 CTCCTTGGGCGCTGCAGAGCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCA 420
|||||
QY 421 GCTCACAAGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGTG 480
|||||
Db 421 GCTCACAAGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGTG 480
|||||
QY 481 CGTTTCTGATGCTGTAGGAGGTCCACCTCTGCGTCAGGCGGGCC 528
|||||
Db 481 CGTTTCTGATGCTGTAGGAGGTCCACCTCTGCGTCAGGCGGGCC 528
|||||

RESULT 4

LOCUS AR087571 1062 bp DNA PAT 07-SEP-2000
DEFINITION Sequence 18 from patent US 5989537.
ACCESSION AR087571
VERSION AR087571.1 GI:10014334
KEYWORDS
SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1062)

AUTHORS Holly,R.D., Lok,S., Foster,D.C., Hagen,F.S., Kaushansky,K.,
Kuijper,J.L., Lofton-Day,C.E. and Oort,P.J.
TITLE Methods for stimulating granulocyte/macrophage lineage using
thrombopoietin
JOURNAL Patent: US 5989537-A 18 23-NOV-1999;
FEATURES Location/Qualifiers
source 1..1062

BASE COUNT 223 a 375 c 236 g 228 t
ORIGIN

Query Match 48.3%; Score 523.2; DB 6; Length 1062;
Best Local Similarity 99.4%; Pred. No. 5e-99;
Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATTGCTCTCGTGTGCTATGCTTCTCTTAAGCTGAAGGCTAACGCTG 60
|||||
Db 1 ATGGAGCTGACTGAATTGCTCTCGTGTGCTATGCTTCTCTTAAGCTGAAGGCTAACGCTG 60
|||||
QY 61 TCCAGCCGGCTCTCGTGTGCTGAGCTCGGAGTCTCAGTAAGTCTGCTGACTCC 120
|||||
Db 61 TCCAGCCGGCTCTCGTGTGCTGAGCTCGGAGTCTCAGTAAGTCTGCTGACTCC 120
|||||
QY 121 CATGCTCTTACAGCAGACTGAGCCAGTGCAGAGGTTACCCCTTTGCCCTACACCTGTC 180
|||||
Db 121 CATGCTCTTACAGCAGACTGAGCCAGTGCAGAGGTTACCCCTTTGCCCTACACCTGTC 180
|||||
QY 181 CTGCTGCTGTGGAGCTTTAGCTTGGGAGATGGAACCCAGATGGAGGAGACCAAG 240
|||||
Db 181 CTGCTGCTGTGGAGCTTTAGCTTGGGAGATGGAACCCAGATGGAGGAGACCAAG 240
|||||
QY 241 GCACAGACATCTGGGAGCAGTCAACCTTCTGCTGGAGGAGTGATGGCAGCAGCGGA 300
|||||
Db 241 GCACAGACATCTGGGAGCAGTCAACCTTCTGCTGGAGGAGTGATGGCAGCAGCGGA 300
|||||
QY 301 CAACTGGGACCCACTTGCCTCTCATCCCTCTGCTGGGAGGCTTTCTGGACAGGTCGCTC 360
|||||
Db 301 CAACTGGGACCCACTTGCCTCTCATCCCTCTGCTGGGAGGCTTTCTGGACAGGTCGCTC 360
|||||
QY 361 CTCCTTGGGCGCTGCAGAGCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCA 420
|||||
Db 361 CTCCTTGGGCGCTGCAGAGCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCA 420
|||||
QY 421 GCTCACAAGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGTG 480
|||||
Db 421 GCTCACAAGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGTG 480
|||||
QY 481 CGTTTCTGATGCTGTAGGAGGTCCACCTCTGCGTCAGGCGGGCC 528
|||||
Db 481 CGTTTCTGATGCTGTAGGAGGTCCACCTCTGCGTCAGGCGGGCC 528
|||||

RESULT 5

LOCUS I49760 1062 bp DNA PAT 07-OCT-1997
DEFINITION Sequence 3 from patent US 5641655.
ACCESSION I49760
VERSION I49760.1 GI:2471980
KEYWORDS
SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1062)

AUTHORS Foster,D.C., Heipel,M.D. and Holly,R.D.
TITLE Methods for producing thrombopoietin polypeptides using a mammalian
tissue plasminogen activator secretory peptide
JOURNAL Patent: US 5641655-A 3 24-JUN-1997;
FEATURES Location/Qualifiers
source 1..1062

6

```

Db 325 CAACCTGGAGCCCACTGGCTCTCATCCCTCCCTGGGGCAGCTTTTCTGGACAGGTCCCTCTC 384
QY 361 CTCTTTGGGGCCCTGAGAGCTCTTGGAAACCCAGCTTCCTCCACAGGGCAGGACCACA 420
Db 385 CTCTTTGGGGCCCTGAGAGCTCTTGGAAACCCAGCTTCCTCCACAGGGCAGGACCACA 444
QY 421 GCTCACAGGATCCCAATGCTTCCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
Db 445 GCTCACAGGATCCCAATGCTTCCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 504
QY 481 CGTTTCTGATGCTTGTAGGAGGCTCCACCTCTGCTGAGGGTGGC 528
Db 505 CGTTTCTGATGCTTGTAGGAGGCTCCACCTCTGCTGAGGGGCC 552

RESULT 9
E12214
LOCUS E12214 1267 bp DNA PAT 24-JUN-1998
DEFINITION Human cDNA encoding thrombopoietin.
ACCESSION E12214
VERSION E12214.1 GI:3251048
KEYWORDS JP 1996291196-A/1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1267)
AUTHORS Miyazaki,H., Kato,H., Ogami,K., Iwamatsu,A., Akahori,H., Kuroki,R., Shinizu,T. and Muto,T.
TITLE PROTEIN HAVING TPO ACTIVITY.
JOURNAL Patent: JP 1996291196-A 1 05-NOV-1996;
COMMENT OS Homo sapiens (human)
PN JP 1996291196-A/1
PD 05-NOV-1996
PF 14-FEB-1995 JP 1995355052
PR 14-FEB-1994 JP 94P 39090, 25-MAR-1994 JP 94P 79842, PR
01-JUN-1994 JP 94P 155126, 15-JUN-1994 JP 94P 167328, PR
17-AUG-1994 JP 94P 227159, 17-AUG-1994 JP 94P 193169, PR
01-NOV-1994 JP 94P 304167, 01-DEC-1994 JP 94P 298669, PR
28-DEC-1994 JP 94P 341200
PI MIYAZAKI HIROSHI, KATO HISASHI, OGAMI KINYA, IWAMATSU AKIHIKO,
PI AKAHORI HIRONORI, KUROKI RYOTA, SHIMIZU TOSHIYUKI, PI MUTO
TAKANORI
PC C07K14/53,C07K1/22,C12N15/09//A61K38/00,C12P21/08; CC
strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FT source 1..1267
FT FT /organism='Homo sapiens'
FT FT /tissue_type='liver'
FT CDS 25..1086
FT FT /product='thrombopoietin'
FT FT sig_peptide 25..87
FT FT mat_peptide 88..1083
FT FT /product='thrombopoietin'.
FEATURES
source 1..1267
/organism='Homo sapiens'
/db_xref='taxon:9606'
BASE COUNT 281 a 436 c 279 g 271 t
ORIGIN

Query Match 48.3%; Score 523.2; DB 6; Length 1267;
Best Local Similarity 99.4%; Pred. No. 4.8e-99;
Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATTGCTCTGCTGGTTCATGCTTCCTTAACGCAAGGCTAACGCTG 60
Db 25 ATGGAGCTGACTGAATTGCTCTGCTGGTTCATGCTTCCTTAACGCAAGGCTAACGCTG 84

```

```

QY 61 TCCAGCCCCGGCTCCTCTGCTTGTGACCTCCGAGTCTCTCAGTAACTGCTTCGTGACTCC 120
Db 85 TCCAGCCCCGGCTCCTCTGCTTGTGACCTCCGAGTCTCTCAGTAACTGCTTCGTGACTCC 144
QY 121 CATGTCTTTCACAGCAGACTGAGCCAGTCCCGAGAGGTTTCAACCTTTGCCCTACACCTGTC 180
Db 145 CATGTCTTTCACAGCAGACTGAGCCAGTCCCGAGAGGTTTCAACCTTTGCCCTACACCTGTC 204
QY 181 CTGCTGCCCTGCTGTGACTTTAGCTTTGGGAGATGGAAACCCAGATGGAGGAGACCAAG 240
Db 205 CTGCTGCCCTGCTGTGACTTTAGCTTTGGGAGATGGAAACCCAGATGGAGGAGACCAAG 264
QY 241 GCACAGGACATTTCTGGGAGCAGTACCCCTTCTGCTGGAGGAGTGTATGGCAGACGGGGA 300
Db 265 GCACAGGACATTTCTGGGAGCAGTACCCCTTCTGCTGGAGGAGTGTATGGCAGACGGGGA 324
QY 301 CAACCTGGGACCCACTTGCCTCTCATCCCTCCCTGGGGCAGCTTTCTGGACAGGTCCGCTCTC 360
Db 325 CAACCTGGGACCCACTTGCCTCTCATCCCTCCCTGGGGCAGCTTTCTGGACAGGTCCGCTCTC 384
QY 361 CTCTTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCACA 420
Db 385 CTCTTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCACA 444
QY 421 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGTCTCCGAGGAAAGGTG 480
Db 445 GCTCACAAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGTCTCCGAGGAAAGGTG 504
QY 481 CGTTTCTGATGCTTGTAGGAGGCTCCACCTCTGCTGAGGGTGGC 528
Db 505 CGTTTCTGATGCTTGTAGGAGGCTCCACCTCTGCTGAGGGGCC 552

RESULT 10
AR008878
LOCUS AR008878 1342 bp DNA PAT 04-DEC-1998
DEFINITION Sequence 1 from patent US 5756083.
ACCESSION AR008878
VERSION AR008878.1 GI:3967683
KEYWORDS UNKNOWN.
SOURCE UNKNOWN.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1342)
AUTHORS Elliott,S.G.
TITLE Mpl ligand analogs
JOURNAL Patent: US 5756083-A 1 26-MAY-1998;
FEATURES Location/Qualifiers
source 1..1342
/organism='unknown'
BASE COUNT 299 a 454 c 293 g 296 t
ORIGIN

Query Match 48.3%; Score 523.2; DB 6; Length 1342;
Best Local Similarity 99.4%; Pred. No. 4.7e-99;
Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATTGCTCTGCTGGTTCATGCTTCCTTAACGCAAGGCTAACGCTG 60
Db 36 ATGGAGCTGACTGAATTGCTCTGCTGGTTCATGCTTCCTTAACGCAAGGCTAACGCTG 95
QY 61 TCCAGCCCCGGCTCCTCTGCTTGTGACCTCCGAGTCCCTCAGTAACTGCTTCGTGACTCC 120
Db 96 TCCAGCCCCGGCTCCTCTGCTTGTGACCTCCGAGTCCCTCAGTAACTGCTTCGTGACTCC 155
QY 121 CATGTCTTTCACAGCAGACTGAGCCAGTCCCGAGAGGTTTCAACCTTTGCCCTACACCTGTC 180
Db 156 CATGTCTTTCACAGCAGACTGAGCCAGTCCCGAGAGGTTTCAACCTTTGCCCTACACCTGTC 215
QY 181 CTGCTGCCCTGCTGGACTTTAGCTTTGGGAGATGGAAACCCAGATGGAGGAGACCAAG 240
Db 216 CTGCTGCCCTGCTGGACTTTAGCTTTGGGAGATGGAAACCCAGATGGAGGAGACCAAG 275

```


Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1691)
Oqami,K.

Submitted (08-JUL-1994) to the DDBJ/EMBL/GenBank databases. Kinya Ogami, KIRIN Brewery Co., Ltd., Pharmaceutical Research Laboratory; 2-2 Soujiamachi 1 chome, Maebashi, Gunma 371, Japan
(Tel: 0272-54-8618; Fax: 0272-52-2307)

2 (slices)
Kato, T., Ogami, K., Shimada, Y., Iwamatsu, A., Sohma, Y., Akahori, H.,
Horie, K., Kokubo, A., Kudo, Y., Maeda, E., Kobayashi, K., Ohashi, H.,
Ozawa, T., Inoue, H., Kawamura, K. and Miyazaki, H.
Purification and characterization of thrombopoietin
J. Biochem. 118 (1), 229-236 (1995)

Erratum: [[published erratum appears in J Biochem (Tokyo) 1996 Jan;119(1):208]]
96015174
J. Biochem. 118 (1): 225-230 (1995)

```

Jan,115(1):206]]
Location/Qualifiers
1..1691
/organism="Homo sapiens"
/db xref="taxon:9606"

```

```

/tissue_type="liver"
102. .164
102. .1163
ptide

```

```
4041:1103
/codon_start=1
/product="thrombopoietin"
/protein_id="BAA21930.1"
/db_xref="GI:2351118"
```

700/analysis = "MELTCELLVMMLLTARUTLSSPAPPACDLRVLSKLLRDSHVILV
SRSCSPVHPVLPVLPVAVDFSGEWKTEETHAKDILGVAITLIEGVNNAARQL
GTCSSLLGSGOVRLLGALSLQVLPQOGRTHAKDPNAIFLSVGLHRKVG
RFLMVGSTGCRAPPTVAVPSISVLINLHNETGTFLETSARTASGSL
LKWOGFRAKTPGLNQTNRSLDQPGVNLHLLNGTRGLFGPSRRRLGCAPOISS
TSDTSGVLPNLCQPSYSPHPPTGQVTLFPLPTLPVTVQLHPLPDPSPAPTPT
TSPILNTSYTHSONLSORE"

	383 a	542 c	371 g	395 t
ISPLENTSIHSQLSQEG				

ch	48.3%;	Score 523.2;	DB 9;	Length 1691;
l Similarity	99.4%;	Pred. NO. 4.4e-99;		
325: Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

TCGAGAGCTGACTGAATTGTCCTTCGTGGTCATGCTTCTCCTAACTGCAAGGCTAACGCTG 60

CGAGCCGGCTCCTCCTGCTTGTGACCTCCGAGTCTCAGTAACTGCTCGTACTCC 120

CCAGCCCGGCTCCTCCTGCTTGTGACCTCCGAGCTCCAGTAACTGCTTCGTGACTCC 221

TTGTCTTTCAGCAGACTGAGCCAGTGCCAGAGGTTACCCCTTTGCCCTACACCTGTC 281

TCTGCTGCTGCTGTGGACTTTAGCTTGGGAGAAATCGAAAACCCAGATGGAGGAGACC**AAG** 240
TCTGCTGCTGCTGTGGACTTTAGCTTGGGAGAAATCGAAAACCCAGATGGAGGAGACC**AAG** 341

ACAGGACATTCTGGGAGCAGTGACCTTCTGCTGGAGGGGACTGATGGCAGCACGGGA 300

ACAGGACATTCTGGGAGCAGTGACCCCTTCTGCTGGAGGGAGTGTATGGCAGCACGGGGA 401

AAC TGGGAGCCACATTGCGCTCTCATCCCTCCTGGGGCAGCTTTCTGGACAGGTCCGTC 360

TACTGGGACCCACATTGGCTCTCATCCTCCTGGGGCAGCTTCTGGACAGGTCCGTC

461

"CCTTGGGCCCCGACAGCCTCCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCACA 420
 |||||
 "CCTTGGGCCCCGACAGCCTCCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCACA 521

TTTCAAGGATCCCAATGCCATCTTCTTGAGCTTCCAACACTGCTCCGAGGAAGGTG 480

RESULT 13
T85129

LOCUS	185129	1342 bp	DNA	PAT	04-APR-1998
DEFINITION	Sequence 1 from patent US 5696250.				
ACCESSION	185129				
VERSION	185129.1	GI:3022649			
KEYWORDS					

SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE
Unknown.				
Unknown.				
Unclassified.				
1 (bases 1 to 1342)				
Elliot, S.G.				
DNA encoding megakaryocyte growth and development factor analogs				

JOURNAL Patent: US 5696250-A 1 09-DEC-1997;
FEATURES Location/Qualifiers
source 1. .1342

	/organism="unknown"		
BASE COUNT	299 a	454 c	293 g
ORIGIN			296 t

Query Match	48.3%;	Score 523.2;	DB 6;	Length 1342;
Best Local Similarity	99.4%;	Pred. No. 4.7e-99;		

	Matches	525;	Conservative	0;	Mismatches	3;	Indels	0;	Gaps	0;
Qy	1	ATGGAGCTGACTGAATTCTCTCTGCTGGTCATGCTTCTCTAACTCGAGGGCTAAGCTTG	60							

36 ATGGAGCTGACCTGAATTGCTCTCTGTTGGTCTACTGCTTCTTAAGCTGAAGCTTAAGCGCTG 95
Db
61 TCAGAGCCCGGCTCCCTGCTGCTTTGAGACTCGAGTCTCTAGTAACTGCTTCGTGACTCC 120
QY
96 TCAGAGCCCGGCTCCCTGCTGCTTTGAGACTCGAGTCTCTAGTAACTGCTTCGTGACTCC 155
Db
121 CATGTCTTTACAGACAGACTCAGAGCGAGTGCCAGAGGTTTCAACCTTTTGCTCAACACTGTG 180
QY

Q7
124 CATGACCTTCACAGCAGACTGAGCCAGTGCCAGAGTTACCCCTTGCCATCACCGTC 180

Db
156 CATGTCTTCACAGCAGACTGAGCCAGTGCCAGAGTTACCCCTTGCCATCACCGTC 215

Dbb

Db 276 GCACAGGACATTCTGGGAGCAGTGACCTTCTGCTGGAGGAGTGATGGCAGCAGCGGGA 335

Qy 301 CAACTGGGACCCACTGGCCCTCATCCCTCTGGGACGTTTCTGGACAGTCCGTC 360
=====

Db 336 CAACTGGGACCCACTGGCCCTCATCCCTCTGGGACGTTTCTGGACAGTCCGTC 395

Qy 361 CTCTTGGGGCCCTGGCAGAGCCTCCTTGGACCCAGCTTCCTCCACAGGGCAGGACCACA 420
|||||

Db	396	CTCCTTGGGGCCCTGCAGAGCCTCCTTGGAAACCCAGTTCCTCCACAGGGGAGGACCACA	455
Oy	421	GCTCACAGGATCCCAATGCGCATCTTCCTGAGCTTCCAACACCTGCTCCGAGGAAGGTG	480

Db 456 GCTCAAGGATCCCAATGCCATCTCTGAGCTTCCACACCTGCTCCGAGGAAGGTG 515

Qy 481 CGTTTCCTGATGTCTGTAGGAGGGTCCACCTCTGCCTCAGGGGTGCC 528
 | | | | | | | | | | | | | | | | | | | | | |
Dd 516 CGTTTCCTGATGTCTGTAGGAGGGTCCACCTCTGCCTCAGCGGSCC 563

RESULT 14
832047

D32047	D32047	1691 bp	PRI	03-SEP-1997
LOCUS				
DEFINITION	Homo sapiens mRNA for thrombopoietin, complete cds.			
ACCESSION	D32047			

ACCESSION D32047
VERSION D32047.1 GI:2351117
KEYWORDS thrombopoietin,
Homo sapiens liver cDNA, cDNA

SOURCE	ORGANISM	CDNA TO mRNA
Homo sapiens liver	Homo sapiens	

||||| 522 GCTCAAGGATCCCAATGCCATCTCTAGCTTCCAAACACCTGCTCCGAGGAAAGTG 581
QY 481 CGTTTCTTGATGCTGTAGGAGGTCACCTCTCGGTGAGGGTGGC 528
Db 582 CGTTTCTGATGCTGTAGGAGGTCACCTCTCGGTGAGGGGCCC 629

RESULT 15
LOCUS E16668 1721 bp DNA PAT 28-JUL-1999
DEFINITION cDNA encoding thrombopoietin.
ACCESSION E16668
VERSION E16668.1 GI:5711351
KEYWORDS JP 1998212243-A/2.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1721)
Kuroda, K.
COMPOSITION CONTAINING THROMBOPOIETIN PROTEIN AND USED FOR
STIMULATING RECOVERY OF HEMOSTATIC THROMBUS-FORMING ABILITY OF
PATENT: JP 1998212243-A 2 11-AUG-1998;
KIRIN BREWERY CO LTD
OS Homo sapiens (human)
PN JP 1998212243-A/2
PD 11-AUG-1998
PF 31-JAN-1997 JP 1997019549
PI KURODA KENJI
PC A61K38/00,A01N1/02,C07K14/52//C12N15/09,C12P21/02,(C12P21/02,
C12R1:19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FH source 1..1721
FT /organism='Homo sapiens'
FT /tissue_type='Liver'
FT 1..101
FT 5'UTR
FT sig_peptide 102..164
FT /product='signal peptide of thrombopoietin' FT
FT mat_peptide 165..1160
FT /product='thrombopoietin'
FT CDS 102..1163
FT /product='thrombopoietin'
FT 1164..1721.
FT 3'UTR
FEATURES
source 1..1721
/organism='Homo sapiens'
/db_xref='taxon:9606'
BASE COUNT 413 a 542 c 371 g 395 t
ORIGIN

Query Match 48.3%; Score 523.2; DB 6; Length 1721;
Best Local Similarity 99.4%; Pred. No. 4.4e-99;
Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATTCCTCGTGGTCATGCTTCTCCTAAGTGAAGGCTAACGCTG 60
Db 102 ATGGAGCTGACTGAATTCCTCGTGGTCATGCTTCTCCTAAGTGAAGGCTAACGCTG 161
QY 61 TCCAGCCCGGCTCCTCGTGGTCATGCTTCTCCTAAGTGAAGGCTAACGCTG 120
Db 162 TCCAGCCCGGCTCCTCGTGGTCATGCTTCTCCTAAGTGAAGGCTAACGCTG 221
QY 121 CATGTCCTTACAGCAGACTGAGCCAGTGCCAGAGGTTACCCCTTTGCCCTACACCTGTC 180
Db 222 CATGTCCTTACAGCAGACTGAGCCAGTGCCAGAGGTTACCCCTTTGCCCTACACCTGTC 281

QY 181 CTGCTGCCTGCTGTGGACTTTAGCTTGGAGAAATGAAAAACCCAGATGGAGGAGACCAAG 240
Db 282 CTGCTGCCTGCTGTGGACTTTAGCTTGGAGAAATGAAAAACCCAGATGGAGGAGACCAAG 341
QY 241 GCACAGGACATTCCTGGGAGCAGTACCCTTCTGTGGAGGAGTGTATGCCAGCAGCGGGA 300
Db 342 GCACAGGACATTCCTGGGAGCAGTACCCTTCTGTGGAGGAGTGTATGCCAGCAGCGGGA 401
QY 301 CAACCTGGACCCACTTGCTCTCATCCCTCTGGGGCAGCTTCTGGACAGTCCGCTC 360
Db 402 CAACCTGGACCCACTTGCTCTCATCCCTCTGGGGCAGCTTCTGGACAGTCCGCTC 461
QY 361 CTCCTTGGGGCCCTGCAGAGCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCAACA 420
Db 462 CTCCTTGGGGCCCTGCAGAGCTCTTGGAAACCCAGCTTCTCCACAGGGCAGGACCAACA 521
QY 421 GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGTG 480
Db 522 GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGTG 581
QY 481 CGTTTCTGATGCTGTAGGAGGTCACCCCTCTCGGTGAGGGTGGC 528
Db 582 CGTTTCTGATGCTGTAGGAGGTCACCCCTCTCGGTGAGGGTGGC 629

Search completed: April 5, 2002, 10:29:34
Job time: 4767 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2002, 10:33:05 ; Search time 188.75 Seconds
(without alignments)
4919.116 Million cell updates/sec

Title: US-09-680-514-6

Perfect score: 1083

Sequence: 1 ATGGAGCTGACTGAATTGCT.....TACGCCACCTTGCCAGCC 1083

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_1101.*
1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1083	100.0	1083	17	AAT41787
2	1030.2	95.1	1095	17	AAT41788
3	993	91.7	1047	17	AAT41786
4	523.2	48.3	861	16	AAQ99552
5	523.2	48.3	1062	16	AAQ99552
6	523.2	48.3	1062	16	AAQ99552
7	523.2	48.3	1062	17	AAQ99552
8	523.2	48.3	1062	17	AAQ99552
9	523.2	48.3	1062	17	AAQ99552
10	523.2	48.3	1062	18	AAQ99555
11	523.2	48.3	1062	19	AAV21696

12	523.2	48.3	1062	21	AA51991	Human thrombopoietin
13	523.2	48.3	1086	17	AAQ47958	Human thrombopoietin
14	523.2	48.3	1342	16	AAQ99704	Human MGDF-1 and M
15	523.2	48.3	1342	17	AAQ36657	Native human mpl 1
16	523.2	48.3	1342	17	AAQ10025	Human MGDF-1/2 cDN
17	523.2	48.3	1721	16	AAQ99554	Thrombopoietin cod
18	523.2	48.3	1721	21	AAZ40191	Human wild type th
19	523.2	48.3	1774	21	AAQ47793	Human thrombopoiet
20	523.2	48.3	1795	16	AAQ94107	hML cDNA. Homo sa
21	523.2	48.3	1795	18	AAQ64318	Human thrombopoiet
22	521.6	48.2	600	17	AAQ36658	Truncated human mp
23	521.6	48.2	605	20	AAQ32813	Human truncated mp
24	521.6	48.2	1267	16	AAQ99553	Thrombopoietin cod
25	515.6	47.6	525	11	AAQ04482	Plasmid PASN6 enco
26	512.4	47.3	525	11	AAQ04484	Plasmid PASN145 en
27	507.6	46.9	525	11	AAQ04481	Plasmid PAS28 enco
28	500.2	46.2	1342	17	AAQ33933	Human megakaryocyt
29	500.2	46.2	1342	19	AAV29068	Nucleotide sequenc
30	500.2	46.2	1342	20	AAQ32812	Human mpl ligand e
31	489.2	45.2	1050	17	AAQ41901	Human thrombopoiet
32	486.8	44.9	525	9	AAQ80947	G-CSF gene isolate
33	486.8	44.9	1520	8	AAQ70223	Plasmid pBRV2 inse
34	486.8	44.9	1520	8	AAQ71320	Sequence encoding
35	486.8	44.9	1521	9	AAQ60937	Plasmid pBRV2 inse
36	486.8	44.9	1525	9	AAQ81478	Sequence encoding
37	486.8	44.9	1525	10	AAQ91086	Plasmid pP12 contg
38	486.6	44.9	1017	16	AAQ97181	PMON15937 DNA enco
39	486.6	44.9	1017	21	AAQ03735	Human interleukin-
40	485.2	44.8	1415	8	AAQ71089	Sequence of human
41	479.8	44.3	741	17	AAQ41907	Human thrombopoiet
42	478	44.1	861	17	AAQ41902	Human thrombopoiet
43	478	44.1	861	17	AAQ15486	Megakaryocyte diff
44	478	44.1	946	17	AAQ29735	Human platelet pro
45	478	44.1	946	21	AAZ44174	Human platelet-pro

ALIGNMENTS

RESULT 1
AAT41787
ID AAT41787 standard; DNA; 1083 BP.
XX
AC AAT41787;
XX
DT 01-JUL-1997 (first entry)
XX
DE Fusion peptide #2 having G-CSF and TPO activity.
XX
KW Fusion protein; human granulocyte colony stimulating factor; hG-CSF;
KW thrombopoietin; TPO; spacer peptide; blood platelet production;
KW leukocyte production; anaemia; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT sig_peptide 1..63
FT mat_peptide 64..1083
FT /*tag= a
FT /*tag= b
XX
PN WO9634016-A1.
XX
PD 31-OCT-1996.
XX
PF 26-APR-1996; 96WO-JP01157.
XX
PR 26-APR-1995; 95JP-0102625.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Anazawa H, Konishi N, Shiotsu Y, Tamaoki T, Terasaki Y;
PI Uchida K, Yamasaki M, Yamashita K, Yokoi H;

XX WPI; 1996-497573/49.
 DR P-PSDB; AAW00378.
 XX
 XX Fusion peptide having G-CSF and with thrombopoietin activity -
 PT optionally chemically modified with a poly:alkylene glycol, used for
 PT treatment of anaemia
 XX
 XX Claim 3; Page 49-51; 71pp; Japanese.
 PS
 CC The sequences given in AAT41786-88 encode fusion proteins which are
 CC composed of a peptide having human granulocyte colony stimulating
 CC factor (hG-CSF) activity fused with a peptide having thrombopoietin
 CC (TPO) activity, opt. via a spacer peptide. Peptides derived from
 CC these by deletion, insertion or substitution of one or more amino
 CC acid residues are included within the scope of the invention. The
 CC fusion peptides stimulate blood platelet and leukocyte production
 CC and are useful in the treatment of anaemia.
 XX
 XX Sequence 1083 BP; 197 A; 355 C; 305 G; 226 T; 0 other;
 SQ

Query Match 100.0%; Score 1083; DB 17; Length 1083;
 Best Local Similarity 100.0%; Pred. No. 2.2e-249; Indels 0; Gaps 0;
 Matches 1083; Conservative 0; Mismatches 0;

QY 1 ATGGAGCTGACTGAATGTCTCGTGGTCATGCTTCTTAACCTCAAGGCTAAGCGTGC 60
 DB 1 atggagctgactgaattgtctcgtgggtcattgtcttcttaactgaaggttaacgctg 60

QY 61 TCACAGCCGCTCTCTGCTGTGTGACCTCCGAGTCCGAGTCTCAGTAAACTGCTTCTGACTTC 120
 DB 61 tcacagccgctctctgctgtgtgacctccgagtcctcagtaactgcttctgacttc 120

QY 121 CATGCTCTTCACAGCAGACTGAGCCAGTGCACGAGTTCACCCCTTTCGCTACACTGCTC 180
 DB 121 catgctcttcacagcagactgagccagtgcacgagttcaccccttgcctacactgtc 180

QY 181 CTGCTGCTCTGTGGAGCTTTAGCTTGGGAGATGAAACCCAGATGAGGAGACCAAG 240
 DB 181 ctgctgctctgtggagcttttagcttgggagatgaaacccagatgagagagaccaag 240

QY 241 GCACAGACATTCGGAGCAGTGGACCTTCTGCTGGAGGAGTGTATGCGACACAGCGGA 300
 DB 241 gcacagacatctggagcagtggacctctctgtgagggagtgtatgagcagcagggga 300

QY 301 CAATGGGACCCACTTGCCTCTCATCCCTCCCTGGGACGCTTTCGGACAGTCCGCTC 360
 DB 301 caatgggacccacttgcctctcatccctccctgggagcgttctcggaggtccgctc 360

QY 361 CTCCTTTGGGCCCCGAGAGCTCTCTTGGAGCCAGCTTCTCCACAGGAGGACCCACA 420
 DB 361 ctcccttggggccccgagagctctcttggagccagcttctccacagggcagaccaca 420

QY 421 GCTCACAAGATCCCAATGCCATCTTCCTGAGCTTCCACACCTGCTCCGAGGAAAGTG 480
 DB 421 gctcacaagatcccaatgccatcttccctgagcttccacacactgctccgaggaagtg 480

QY 481 CGTTTCTGATGCTCTAGAGGGTCCACCCCTCTCGCTCAGGGTGGCGGTTCTGGAGGT 540
 DB 481 cgtttctgatgctctagaggggtccacccctctcgtcaggggtggcggttctggaggt 540

QY 541 GGTTCGGAGGGGGTCTAGAGCACCACATATCGCGCTCCAGTCTACACAGAGCTTC 600
 DB 541 ggttcgaggggggtctagagcaccacacatatcgcgctcagtcctaccacagagcttc 600

QY 601 CTTTTTAAAAGCTTAGAGCAAGTAGGAAAGATCCAGGGCGATGGCGAGCGCTCCAGGAG 660
 DB 601 ctttttaaaagcttagagcaagttaggaagatccagggcgatggcgagcgctccagag 660

QY 661 AAGCTGTGTGCCACCTACAGCTGTGCCACCCCGAGGAGCTGGTCTGCTGGACACTCT 720
 DB 661 aagctgtgtgccacctacagctgtgccaccccgagagctgggtctgctgagacactct 720

QY 721 CTGGGCATCCCTGGGCTCCCTGAGCAGCTGCCAGCCAGCCAGCTGCGAGCTGGCAGGC 780
 DB 721 ctgggcataccctgggctccctgagcagctgccagccagccagctgagctggcaggc 780

QY 781 TGCTTGAGCAACTCATACAGCGCTTTTCTTCTTACACAGGGGCTCTGAGGCGCTGGAA 840
 DB 781 tgcttgagcaactcatacagcgcttttcttcttctacacaggggctctgagggccctgaa 840

QY 841 GGGATCTCCCGGAGTTGGTCCACCTTGGACACTGCGAGCTGGACGTGCGGACTTT 900
 DB 841 gggatctcccgagttgggtccaccttggacactgacactgacgtgacgtcgcgacttt 900

QY 901 GCACACCACTCTGGCAGCAGATGGAAGAACTGGGAATGCCCTCTGCCCTGCGAGCCAC 960
 DB 901 gcacaccactctggcagcagatggaagaactgggaatggccctctgagcagccacc 960

QY 961 CAGGGTGCCATCGCGGCTTCGCTCTGCTTTCACAGCGGCGCAGAGGGGCTCTAGTT 1020
 DB 961 cagggtgccatcgcggcttcgctctgcttctccagcgccgaggggtcctagt 1020

QY 1021 GCCTCCCATCTGCAGAGCTTCTGGAGGTGCTGACCGGCTTCTAGCCACCTTGCCAG 1080
 DB 1021 gcctcccatctgcagagcttctggaggtgctgacccggttctacgcccacttgccag 1080

QY 1081 CCC 1083
 DB 1081 ccc 1083

RESULT 2
 AAT41788
 ID AAT41788 standard; DNA; 1095 BP.
 XX
 AC AAT41788;
 XX
 DT 01-JUL-1997 (first entry)
 XX
 DE Fusion peptide #3 having G-CSF and TPO activity.
 XX
 KW Fusion protein; human granulocyte colony stimulating factor; hG-CSF;
 KW thrombopoietin; TPO; spacer peptide; blood platelet production;
 KW leukocyte production; anaemia; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT sig_peptide 1..63
 FT mat_peptide 64..1095
 FT mat_peptide /*tag= a
 FT mat_peptide /*tag= b
 XX
 PN WO9634016-A1.
 XX
 PD 31-OCT-1996.
 XX
 PF 26-APR-1996; 96WO-JP01157.
 XX
 PR 26-APR-1995; 95JP-0102625.
 XX
 PA (KYOW) KYOWA HAKKO KOGYO KK.
 XX
 PI Anazawa H, Konishi N, Shiotsu Y, Tamaoki T, Terasaki Y;
 PI Uchida K, Yamasaki M, Yamashita K, Yokoi H;
 XX
 XX WPI; 1996-497573/49.
 DR P-PSDB; AAW00379.
 XX
 PT Fusion peptide having G-CSF and with thrombopoietin activity -
 PT optionally chemically modified with a poly:alkylene glycol, used for
 PT treatment of anaemia
 XX
 PS Claim 3; Page 52-54; 71pp; Japanese.

XX The sequences given in AAT41786-88 encode fusion proteins which are
CC composed of a peptide having human granulocyte colony stimulating
CC factor (hG-CSF) activity fused with a peptide having thrombopoietin
CC (TPO) activity, opt. via a spacer peptide. Peptides derived from
CC these by deletion, insertion or substitution of one or more amino
CC acid residues are included within the scope of the invention. The
CC fusion peptides stimulate blood platelet and leukocyte production
CC and are useful in the treatment of anaemia.
XX
SQ Sequence 1095 BP; 194 A; 362 C; 311 G; 228 T; 0 other;

Query Match 95.1%; Score 1030.2; DB 17; Length 1095;
Best Local Similarity 97.7%; Pred. No. 9e-237;
Matches 1070; Conservative 0; Mismatches 13; Indels 12; Gaps 2;

QY 1 ATGAGCTGACTGAATGCTCTCGTGGTCATGCTTCTCTAACTGAAGCTTAAGCTG 60
DB 1 atgagctgactgaattgctctcggtggtcatgctctctcctaactgaaggctaaagctg 60

QY 61 TCCAGCCCGGCTCTCTCTGCTTGTGACCTCCGAGTCTCAGTAAGTCTCTGCTGACTCC 120
DB 1 tccagcccggtctctctgcttgtgacctccgagtcctcagtaaaactgctctgactcc 120

QY 121 CATGCTCTTACACAGCAGTGAAGCCAGTGCACAGAGTTCACCCCTTTGCCATACCTGTC 180
DB 121 catgctcttcaagcagactgagcagtgcccaagagttcacaccttgcctacacctgtc 180

QY 181 CTGCTGCCGCTGCTGGACTTTAGCTTGGGAGATGGAACACCCAGATGGAGAGACCAAG 240
DB 181 ctgctgccgctgctggactttagcttgggagatggaacacccagatggagagaccaag 240

QY 241 GCACAGGACATTTGGGACAGTGCACCTTCTGCTGGAGGAGTGTGCGACGACGGGA 300
DB 241 gcacaggacatttgggagcagtgacacctctctgctggagagtgatgacagcacggga 300

QY 301 CAACTGGGACCCACTTGTCTCTCATCCCTCTGCGGAGCTTCTTGACAGGTCCTGCTC 360
DB 301 caactgggacccaacttgtctctcatccctctgctgggagcttcttgacaggtccgtctc 360

QY 361 CTCTTGGGGCCCTGACAGAGCTCTCTTGAACCCAGCTTCTCCACAGGCGACGACACA 420
DB 361 ctcttggggccctgacagagctctcttgaacccagcttctccacagggcaggaacaca 420

QY 421 GCTCACAGGATCCCAATGCATCTTCCCTGAGCTTCCACACCTGCTCCGAGCAAGGTG 480
DB 421 gctcacaggatcccaatgcatttccctgagcttcccaacacctgtctccgagaaagggtg 480

QY 481 CGTTTCTGATGCTTAGGAGGCTCCACCTCTGCGTCAAGG---GGTGGCGGTTCTGGA 537
DB 481 cgttctctgactgttagagaggtccacctctgctacggtcccgaggtggtgctctgctg 540

QY 538 GGTGGTTCGGAGGGGTTTC-----TAGAGCACCAACATATATCGCGCTCGAGTCTA 598
DB 541 ggtggttctgttggtggctccggaggcgtgctgctgacacatctgcgctcagtgatcta 600

QY 589 CCACAGAGCTTCTTTTAAAGCTTAGACCAAGTCCAGGAGATCCAGGGCGATGGCGCA 648
DB 601 ccacagagcttcttttaaagcttagagcaagttagaagatccagagcgatggtgcga 660

QY 649 GCCTCCAGGAGAAGCTGTGTGCCACTACAAAGCTGTGCGACCCCGAGGAGCTGTGCTG 708
DB 661 gcgctccaggagaagctgtgtgccacctacaagctgtgccacctccagagagctgtgctg 720

QY 709 CTCGGACACTCTTGGGATTCCTCTGGGCTGCCCTGAGCAGTGTGCCCGAGGCGCTG 768
DB 721 ctgggacactctctgggactccctgggtccctgagcagctgccccagcagggccctg 780

QY 769 CAGCTGGAGGCTGTGAGCCCAACTCCATAGGGGCTTTTCTCTACAGGGGCTCCG 828
DB 781 cagctggagggctgtgagcccaactcccatagcggccttctctaccaggggctcctg 840

QY 829 CAGGCCCTGGAAGGGATCTCCCCGAGTTGGTCCACCTTGACACACTGCGAGCTGGAC 888
DB 841 caggccctggaaggatctccccgagttgggtccccaccttgacacactgcaagctggac 900

QY 889 GTCCGCGACTTTGCCACCACCATCTGCGACGACATGGAAGAACTGGGAATGGCCCTGCC 948
DB 901 gtccgcgactttgccaccacctctgacagcagatggaagaactgggaatggccctgcc 960

QY 949 CTCAGCCCCACCCAGGGTGCATGCCGGCTTCCCTCTGCTTTCCAGCGCCGGCAGGA 1008
DB 961 ctgagccccacccagggtgccatgcccgtccctgctctctcttccagcgccggcgagga 1020

QY 1009 GGGTCTCTAGTTGCCCTCCCATCTGCAGAGCTTCTCTGAGAGTGTGCTACCGCTTCTACGC 1068
DB 1021 ggggtctctagttgccctcccatctgcagagcttctctgaggtgctgacgcggtctcagc 1080

QY 1069 CACCTTGGCCAGGCC 1083
DB 1081 caccttggccagccc 1095

RESULT 3
AAT41786
ID AAT41786 standard; DNA; 1047 BP.
XX
AC AAT41786;
XX
DT 01-JUL-1997 (first entry)
XX
DE Fusion peptide #1 having G-CSF and TPO activity.
XX
KW Fusion protein; human granulocyte colony stimulating factor; hG-CSF;
KW thrombopoietin; TPO; spacer peptide; blood platelet production;
KW leukocyte production; anaemia; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT sig_peptide 1..63
FT /*tag= a
FT mat_peptide 64..1047
FT /*tag= b
XX
PN W09634016-A1.
XX
PD 31-OCT-1996.
XX
PF 26-APR-1996; 96WO-JP01157.
XX
PR 26-APR-1995; 95JP-0102625.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Anazawa H.; Konishi N, Shiotsu Y, Tamaoki T, Terasaki Y;
PI Uchida K, Yamasaki M, Yamashita K, Yokoi H;
XX
WPI: 1996-497573/49.
DR P-PSDB; AAW00377.
XX
PT Fusion peptide having G-CSF and with thrombopoietin activity -
PT optionally chemically modified with a poly:alkylene glycol, used for
PT treatment of anaemia
XX
PS Claim 3; Page 46-48; 71pp; Japanese.
XX
CC The sequences given in AAT41786-88 encode fusion proteins which are
CC composed of a peptide having human granulocyte colony stimulating
CC factor (hG-CSF) activity fused with a peptide having thrombopoietin
CC (TPO) activity, opt. via a spacer peptide. Peptides derived from
CC these by deletion, insertion or substitution of one or more amino
CC acid residues are included within the scope of the invention. The
CC fusion peptides stimulate blood platelet and leukocyte production
CC and are useful in the treatment of anaemia.
CC

XX This sequence represents the coding sequence for a thrombopoietin (TPO)
 CC protein. TPO is a humoral factor capable of promoting platelet
 CC production. The encoded protein can be purified by using a gel affinity
 CC column where Mpl has been coupled to a resin. AAQ99551 and AAQ99553
 CC also encode TPO proteins. The DNA sequences are inserted into vectors
 CC which are used to transform prokaryotic and eukaryotic host cells. Using
 CC the proteins encoded by these sequences, and derivatives of them,
 CC antibodies specifically immunoreactive with a TPO protein can be created.
 CC The antibodies, DNA sequences and vectors are used to isolate the
 CC protein sequences. The TPO proteins can then be used in the treatment
 CC of platelet disorders. These include thrombocytopenia, hypoplastic
 CC anaemia, AIDS, disseminated intravascular coagulation syndrome and
 CC thrombotic thrombocytopenia.
 XX
 SQ Sequence 861 BP; 243 A; 257 C; 197 G; 164 T; 0 other;

Query Match 48.3%; Score 523.2; DB 16; Length 861;
 Best Local Similarity 99.4%; Pred. No. 9.3e-116;
 Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGAGCTGACTGAATTCGCTCGGTGATGCTTCTCTCACTGCAAGGCTAACGCTG 60
 DB 25 atggagctgactgaattgctcctcggtgctatgcttctcttaactgaagctaaagctg 84
 QY 61 TCAGACCCGGCTCTCTGCTTGTGACCTCGAGTCCAGTCACTCACTGCTTCTGCTGCTC 120
 DB 85 tcagacccggctctctgcttgtgacctcgcagtcctcagtaactgcttctgactcc 144
 QY 121 CATGTCTTCCACAGCAGCTGAGCCAGTGCCTGAGGCTTCCACCTTTGCTTACACCTGTC 180
 DB 145 catgtcttccacagcagactgagccagtgccagaggttcacctttgctacactgtc 204
 QY 181 CTGCTGCTCTGTGGACTTTAGCTGGGAGATGGAACCCAGATGGAGGACCAAG 240
 DB 205 ctgctgctctgtggacttttagcttgggagatggaaacccagatggagagaccaag 264
 QY 241 GCACAGACATTCCTGGGAGCAGTGACCTTCTGCTGGAGGAGTGTGAGCAGCAGGGGA 300
 DB 265 gcacagacattctgggagcagtgacctctctgctggagggagtgatggcagcagcggga 324
 QY 301 CAATGGGACCCACTTGTGCTCTCATCCCTCTGCTGGGCGAGCTTCTGGACAGGTCCTGTC 360
 DB 325 caactgggacccactgtctctcatccctctgctgggagcagcttctggacaggtccgctc 384
 QY 361 CTCTTGGGGCCCTGACAGACCTCTTGGAGCCAGCTTCTCCACAGGCGAGGACACA 420
 DB 385 ctcttggggccctgcagagcctctcttggaaacccagcttctccacagggcagggaccaca 444
 QY 421 GCTCACAAAGGATCCCAATGCCATCTTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
 DB 445 gctcacaaagatcccaatgccattctctgagcttccaaacactgtctccgaggaagggtg 504
 QY 481 CGTTTCTGATGTTGTAGAGGGTCCACCTCTGCTGCTCAGGGGTGGC 528
 DB 505 cgtttctgatgtgttagaggggtccaacctctgctgctcagcgggcc 552

RESULT 5
 AAT03941
 ID AAT03941 standard; cDNA; 1062 BP.

XX
 AC AAT03941;
 XX
 DT 28-APR-1996 (first entry)
 XX Human thrombopoietin coding sequence.
 DE
 DE Thrombopoietin; erythropoiesis stimulator; treatment;
 KW Thrombocytopenia; anaemia; ds.
 KW
 XX Homo sapiens.

XX Key Location/Qualifiers
 FH 1..1062
 FT /*tag= a
 XX
 PN WO9521626-A1.
 XX 17-AUG-1995.
 PD
 XX 09-FEB-1995; 95WO-US01829.
 XX 01-DEC-1994; 94US-0347748.
 PR 14-FEB-1994; 94US-0196025.
 PR 25-FEB-1994; 94US-0203197.
 PR 21-MAR-1994; 94US-0215203.
 PR 01-JUN-1994; 94US-0252491.
 PR 09-AUG-1994; 94US-0288417.
 PR 07-NOV-1994; 94US-0335566.
 XX (UNIW) UNIV WASHINGTON.
 PA
 XX Kaushansky K;
 PI
 XX WPI; 1995-292944/38.
 DR P-PSDB; AAR82682.
 DR
 XX Stimulation of erythropoiesis using thrombopoietin and opt.
 PT erythropoietin - for the treatment of thrombocytopenia and anaemia.
 XX
 PS Disclosure; Page 38-40; 66pp; English.
 XX
 CC This sequence encodes human thrombopoietin, which stimulates
 CC erythropoiesis to produce an increase in proliferation or
 CC differentiation of erythroid cells or to increase reticulocyte
 CC counts at least 2-fold over baseline reticulocyte counts and,
 CC optionally, platelet levels to at least 20000/cu mm. The protein
 CC can be used in a composition, optionally with erythropoietin, for
 CC use in the treatment of thrombocytopenia and anaemia, such as
 CC that caused by destruction of haematopoietic cells in bone marrow,
 CC in the treatment of cancer with chemotherapy and radiation, and in
 CC pathological conditions such as myelodysplasia, AIDS, aplastic
 CC anaemia, autoimmune disease or inflammatory disease.
 CC
 XX Sequence 1062 BP; 223 A; 375 C; 236 G; 228 T; 0 other;

Query Match 48.3%; Score 523.2; DB 16; Length 1062;
 Best Local Similarity 99.4%; Pred. No. 9.8e-116;
 Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGAGCTGACTGAATTCGCTCGGTGATGCTTCTCACTGCAAGGCTAACGCTG 60
 DB 1 atggagctgactgaattgctcctcggtgctatgcttctcttaactgaaggtcaacgctg 60
 QY 61 TCAGACCCGGCTCTCTGCTTGTGACCTCCGAGTCCCTAGTAACTGCTTCTGCTGCTC 120
 DB 61 tcagacccggctctctgcttgtgacctccgagtcctcagtaactgcttctgactcc 120
 QY 121 CATGTCTTCCACAGCAGCTGAGCCAGTGCCTGAGGCTTCCACCTTTGCTTACACCTGTC 180
 DB 121 catgtcttccacagcagactgagccagtgccagaggttcacctttgctacactgtc 180
 QY 181 CTGCTGCTGCTGCTGAGCTTTAGCTTGGGAGATGGAACCCAGATGGAGGACCAAG 240
 DB 181 ctgctgcttctgtgagcttttagcttgggagaaatggaaacccagatggagagaccaag 240
 QY 241 GCACAGACATTCCTGGGAGCAGTGACCTTCTGCTGAGGAGTGTGAGCAGCAGGGGA 300
 DB 241 gcacagacattctgggagcagtgacctctctgctgagggagtgatggcagcagcggga 300
 QY 301 CAATGGGACCCACTTGTGCTCTCATCCCTCTGCTGGGCGAGCTTCTGGACAGGTCCTGTC 360
 DB 301 caactgggacccactgtctctcatccctctgctgggagcagcttcttgagacaggtccgctc 360

```

QY 361 CTCCTTGGGGCCCTGCAGAGCCCTTGGAAACCCAGCTTCTCCACAGGCGAGACCACA 420
Db 361 CTCCTTGGGGCCCTGCAGAGCCCTTGGAAACCCAGCTTCTCCACAGGCGAGACCACA 420
QY 421 GCTCACAGAGTCCCAATGCCATCTTCTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
Db 421 GCTCACAGAGTCCCAATGCCATCTTCTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
QY 481 CGTTTCTGATGCTGTGAGGAGGTCCACCCCTCTGCTGCTGAGGGGTGSC 528
Db 481 CGTTTCTGATGCTGTGAGGAGGTCCACCCCTCTGCTGCTGAGGGGTGSC 528

RESULT 6
AAT04050
ID AAT04050 standard; cDNA; 1062 BP.
XX
AC AAT04050;
DT 28-MAR-1996 (first entry)
DE Sequence encoding haematopoietic protein.
KW Haematopoiesis; protein; bone marrow; stem cells; precursor;
KW platelet; therapy; ds.
XX
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 1..1062
FT FT /*tag= a
FT FT /product= Haematopoietic protein.
PN WO9521920-A1.
PD 17-AUG-1995.
XX
XX 05-AUG-1994; 94WO-US08806.
PR 01-JUN-1994; 94US-0252491.
PR 14-FEB-1994; 94US-0196025.
PR 23-FEB-1994; 94US-0203197.
PR 21-MAR-1994; 94US-0215203.
XX
XX (UNIW ) UNIV WASHINGTON.
XX (ZYMO ) ZYMOGENETICS INC.
XX
PI Burkhead SK, Foster DC, Hagen FS, Holly RD, Kaushansky K;
PI Kuijper JL, Lofton-day C, Lok S, Oort PJ;
XX
XX WPI: 1995-293121/38.
XX P-PSDB; AAR79906.
XX
XX Haematopoietic proteins and polypeptide(s) - useful for in vivo and
XX ex vivo therapy
XX
XX Claim 20; Page 106-108; 137pp; English.
XX
XX Haematopoietic proteins and polypeptides are useful for stimulating
XX platelet production in a mammal when given in a therapeutically
XX effective amount. They are also useful for stimulating bone marrow
XX cell proliferation where the bone marrow cells are megakaryocytes or
XX their precursors.
XX
XX Sequence 1062 BP; 223 A; 375 C; 236 G; 228 T; 0 other;

Query Match 48.3%; Score 523.2; DB 16; Length 1062;
Best Local Similarity 99.4%; Pred. No. 9.8e-116;
Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATGCTCTCTCTGCTGGTCATGCTTCTCTCACTGAAGGCTAACGCTG 60

```

```

Db 1 ATGGAGCTGACTGAATGCTCTCTCTCTGCTGGTCATGCTTCTCTCACTGAAGGCTAACGCTG 60
QY 61 TCACAGCCCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
Db 61 TCACAGCCCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
QY 121 CATGTCTCTTACACAGCAGACTGAGCCAGTGCACAGAGTTCACCCCTTTTGCCTTACACCTGTC 180
Db 121 CATGTCTCTTACACAGCAGACTGAGCCAGTGCACAGAGTTCACCCCTTTTGCCTTACACCTGTC 180
QY 181 CTGTCTCTCTGCTGTGACCTTTAGCTTTGGGAGATGAAACCCAGATGGAGGAGACCAAG 240
Db 181 CTGTCTCTCTGCTGTGACCTTTAGCTTTGGGAGATGAAACCCAGATGGAGGAGACCAAG 240
QY 241 GCACAGGACATTTCTGGGAGCAGTACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
Db 241 GCACAGGACATTTCTGGGAGCAGTACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
QY 301 CAACTGGGACCCACTTGCCTCTCATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
Db 301 CAACTGGGACCCACTTGCCTCTCATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
QY 361 CTCTCTTGGGGCCCTGCAGAGCCTCTCTTGGAAACCCAGCTTCTCTCTCTCTCTCTCTCTCTCTCT 420
Db 361 CTCTCTTGGGGCCCTGCAGAGCCTCTCTTGGAAACCCAGCTTCTCTCTCTCTCTCTCTCTCTCTCT 420
QY 421 GCTCACAGGATCCCAATGCCATCTTCTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
Db 421 GCTCACAGGATCCCAATGCCATCTTCTCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
QY 481 CGTTTCTCTGATGCTGTGAGGAGGTCCACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 528
Db 481 CGTTTCTCTGATGCTGTGAGGAGGTCCACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 528

RESULT 7
AAT34852
ID AAT34852 standard; cDNA; 1062 BP.
XX
AC AAT34852;
XX
XX 03-DEC-1996 (first entry)
XX Human thrombopoietin cDNA.
XX
XX Thrombopoietin; TPO; thrombocytopenia; platelet; therapy; ds.
XX
XX Homo sapiens.
XX
XX WO9617062-A1.
XX
XX 06-JUN-1996.
XX
XX 15-NOV-1995; 95WO-US14929.
XX
XX 30-NOV-1994; 94US-0346999.
XX
XX (ZYMO ) ZYMOGENETICS INC.
XX
XX Buddle MM, Downey W, Forstrom JW, Foster DC, Lok S;
XX Osborn SG;
XX
XX WPI: 1996-286830/29.
XX P-PSDB; AAR98947.
XX
XX New low mol. wt. forms of thrombopoietin and related DNA - vectors
XX and transformed cells, useful for stimulating platelet prodn. esp.
XX for treatment of thrombocytopenia
XX
XX Claim 19; Page 57-59; 84pp; English.
XX
XX A cDNA clone (AAT34852) codes for human thrombopoietin (AAR98947), a

```



```
RESULT 9
AAT32591
ID AAT32591 standard; cDNA; 1062 BP.
XX
XX AC AAT32591;
XX
XX 30-SEP-1996 (first entry)
XX
XX DE Human thrombopoietin cDNA.
XX
XX KW Thrombopoietin; TPO; tissue plasminogen activator; cytokine;
XX KW protein secretion; signal peptide; thrombocytopenia; therapy; ds.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX sig_peptide 1..63
XX FT /*tag= a
XX FT 64..1059
XX FT mat_peptide /*tag= b
XX
XX PN WO9617067-A1.
XX
XX PD 06-JUN-1996.
XX
XX PF 15-NOV-1995; 95WO-US14932.
XX
XX PR 30-NOV-1994; 94US-0347029.
XX
XX PA (ZYMO ) ZYMOGENETICS INC.
XX
XX PI Foster DC, Heipel MD, Holly RD;
XX
XX DR WPI; 1996-277784/28.
XX DR P-PSDB; AAR99599.
XX
XX PT DNA construct encoding native mammalian tissue plasminogen activator
XX PT secretory peptide and thrombopoietin polypeptide - used to produce
XX PT thrombopoietin polypeptide(s) on a large scale and in cost effective
XX PT manner
XX
XX PS Disclosure; Page 41-43; 56pp; English.
XX
XX CC A cDNA clone (AAT32591) codes for human thrombopoietin (TPO) (AAR99599).
XX CC Novel DNA constructs code for a secretory peptide (AAR9600) fused to
XX CC the TPO polypeptide, with a cleavage site at the junction. Such
XX CC constructs can be expressed in transformed host cells, pref.
XX CC Pichia pastoris, Saccharomyces cerevisiae or rodent kidney cells,
XX CC for large-scale, cost-effective prodn. of human TPO, useful for the
XX CC treatment of thrombocytopenia.
XX
XX SQ Sequence 1062 BP; 223 A; 375 C; 236 G; 228 T; 0 other;

Query Match 48.3%; Score 523.2; DB 17; Length 1062;
Best Local Similarity 99.4%; Pred. No. 9.8e-116;
Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGGAGCTGACTGAATGCTCTGCTGGTCAATGCTCTCTTAAGTCAAGGCTAAGCGTG 60
Db 1 atggagctgactgaattgctctctggtcatgcttctcttaactgcaaggctaaacgtg 60
Qy 61 TCAGCCCGGCTCTCTGCTGCTGACCTCCGAGTCTCAGTAACTGCTGCTGACTCC 120
Db 61 tcagcccggtctctctgctgctgacccctccagtcctcagtaaaactgcttctgactcc 120
Qy 121 CATGCTCTTCACAGCAGACTGACCCAGTGCACAGAGGTTTCACCCCTTGTGCTACACCTGTC 180
Db 121 catgctcttcacagcagactgacccagtgccccagaggttcaccccttgcctacacctgtc 180
Qy 181 CTGCTGCTCTCTGACTTTAGCTTGGGAGATGGAACCCAGATGGAGGAGACCAAG 240
Db 181 ctgctgctctctgacttttagcttgggagatggaaaccagatggagagaccaag 240

Db 181 ctgctgctctgctgacttttagcttggagaaatgaaaccagatggagagaccaag 240
Qy 241 GCACAGGACATTTCTGGGAGCAGTACCCTTCTGCTGGAGGGAGTGTATGGCAGACGGGA 300
Db 241 gcacaggacattctggagcagtgacccttctgctggagggagtgatggcagcacggga 300
Qy 301 CAACCTGGGACCCACTTGCTCTCATCCCTCCCTGGGCGAGCTTCTGCTGACAGGTCGCTTC 360
Db 301 caactgggaccacttgctctcatccctccctcctgggagcagcttctgagcaggcccgtctc 360
Qy 361 CTCTCTGGGGCCCTGCAGAGCCTCTTGAACCCAGCTTCTCCACAGGGCGAGGACACACA 420
Db 361 ctctctggggccctgcagagcctctcttgaaacccagctctccacagggcagagaccaca 420
Qy 421 GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCCACACACTGCTCCGAGGAAAGGTG 480
Db 421 gctcaagagatcccaatgccatctctgagcttccacacctgctccgaggaagggtg 480
Qy 481 CGTTTCTGATGCTTGTAGGAGGGTCCACCCCTGCTGCGTCAGGGGTGGC 528
Db 481 cgtttctgactgcttgttaggagggtccaccctctgctcagggcgggcc 528

RESULT 10
AAT85555
ID AAT85555 standard; cDNA; 1062 BP.
XX
XX AC AAT85555;
XX
XX DT 05-NOV-1997 (first entry)
XX
XX DE Human thrombopoietin cDNA.
XX
XX KW TPO; cancer; myeloid cell proliferation; expansion; bone marrow;
XX KW peripheral blood stem cell; chemotherapy; radiation therapy;
XX KW breast cancer; leukaemia; lymphoma; multiple myeloma; ds.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX CDS 1..1062
XX FT /*tag= a
XX FT 64..1059
XX FT mat_peptide /*tag= b
XX FT /product= Thrombopoietin
XX
XX PN WO9640218-A1.
XX
XX PD 19-DEC-1996.
XX
XX PF 29-MAY-1996; 96WO-US07880.
XX
XX PR 07-JUN-1995; 95US-0482212.
XX
XX PA (ZYMO ) ZYMOGENETICS INC.
XX
XX PI Fibbe WE, Grossmann A;
XX
XX DR WPI; 1997-108626/10.
XX DR P-PSDB; AAW26841.
XX
XX PT Increasing haematopoietic cells in patient - by admin. of stem cells
XX PT from donor previously treated with thrombopoietin, used in patients
XX PT who have received chemotherapy or radiation for cancer
XX
XX PS Disclosure; Page 17-19; 32pp; English.
XX
XX CC In a claimed method, thrombopoietin (TPO) (especially human TPO) is
XX CC administered to a donor to stimulate proliferation of myeloid cells.
XX CC Bone marrow or peripheral blood stem cells are collected from the
XX CC donor and administered to the recipient. The method is used for
XX CC preparing cells for (bone marrow) transplantation and for stimulating
```

CC platelet or erythrocyte recovery in a patient receiving chemotherapy
CC or radiation therapy for e.g. breast cancer, leukaemia, lymphoma or
CC multiple myeloma. Treating the donor with TPO accelerates restoration
CC of erythrocyte and thrombocyte levels in the patient after transplant.
CC This restoration may be further improved by administering TPO to the
CC recipient after transplant. The present sequence encodes human TPO.
XX
SQ Sequence 1062 BP; 223 A; 375 C; 236 G; 228 T; 0 other;

Query Match 48.3%; Score 523.2; DB 18; Length 1062;
Best Local Similarity 99.4%; Pred. No. 9.8e-116;
Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATTCCTCCTCGTGTGCTATGCTTCCTCACTGCAAGCTCAACGCTG 60
DB 1 atggagctgactgaattgctcctcctcggtgctatgcttccctcaactgcaagctcaacgctg 60
QY 61 TCCAGCCCGGCTCCTCGTGTGACCTCCGAGTCCCTCAGTAACTGCTTCGTGACTCC 120
DB 61 tccagcccggtcctcctcggtgacccctcgagctccgaggttcacactgcttcgactcc 120
QY 121 CATGCTTTCACAGACTGAGGAGTGCAGAGTGCAGAGTTCACCTTTGCCCTACACCTGTC 180
DB 121 catgcttctcacagactgagcagtgccagaggttcacacttgcctacactgtc 180
QY 181 CTGCTGCTGCTGTGGACTTTAGCTTGGAGATGGAATGGAATGGAATGGAATGGAATGGA 240
DB 181 ctgctgctgctgtggacttttagcttggagatggagatggagatggagatggagatggag 240
QY 241 GCACAGGACATTCCTGGAGCAGTGAACCTTCTGCTGGAGGAGTGAATGGAATGGAATGGA 300
DB 241 gcacaggacattctggagcagtgacccttctgctggagagtgatggagcagcgaggga 300
QY 301 CAATGGGACCACTTGCTCTCATCCCTCTGCTGGGAGCTTTCTGGACAGGTCGCTGTC 360
DB 301 caactgggaccacttgctctcatccctctgctgggagctttctggacaggtcgtctc 360
QY 361 CTCTTGGGGCCCTGCAGAGCTCTTGGAGACCTCTTGGAGACCTCTTCCACAGGGCAGACCA 420
DB 361 ctcttggggccctgcagagctcttggagacctcttggagacctcttccacagggcagaccaca 420
QY 421 GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGTG 480
DB 421 gctcacaggatcccaatgccatcttctgagcttccaaacacctgctccgaggaagtg 480
QY 481 CGTTTCCGTGCTGTGAGGAGGTGCCACCTCTGCGTCAAGGGTGGC 528
DB 481 cgtttccgtgctgtgagaggtgccacctctgctgctcagggcgcc 528

RESULT 11
AAV21696
ID AAV21696 standard; cDNA; 1062 BP.
XX
AC AAV21696;
XX
XX
DT 07-AUG-1998 (first entry)
XX
DE Human thrombopoietin encoding cDNA.
XX
KW Thrombopoietin; TPO; expression vector; platelet number; cytokine;
KW human; bone marrow; proliferation; treatment; cancer; aplastic anaemia;
KW myelodysplastic syndrome; chemotherapy; cytopenia; thrombocytopenia;
KW haematologic disorder; leukaemia; lymphoma; ss.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1062
FT FT /*tag= a
FT FT /product= "Thrombopoietin polypeptide"
FT FT 1..63

FT mat_peptide /*tag= b
FT 64..1059 /*tag= c
XX
PN W09806849-A1.
XX 19-FEB-1998.
XX 30-JUL-1997; 97WO-US13543.
XX 13-AUG-1996; 96US-0696447.
XX (ZYMO) ZYMOGENETICS INC.
XX Irani M, Morrison-nelson GR;
XX WPT; 1998-159541/14.
XX P-PSDB; AAW53124.
XX
XX Thrombopoietin protein expression vector - used for increasing
XX platelet number in a mammal
XX
XX Disclosure; Pages 29-32; 56pp; English.
XX
XX This cDNA encodes a thrombopoietin (TPO) polypeptide. This can be used
XX in the construction of a new expression vector replicable in a
XX eukaryotic host cell encoding TPO polypeptides. The vector comprises a
XX transcription promoter, a first DNA segment encoding a secretory leader,
XX a second segment encoding a TPO polypeptide and a transcription
XX terminator which are all operably linked. The second TPO segment consists
XX of C-X-B, where C is a human TPO cytokine domain, X is a peptide bond or
XX a linker consisting of one or two amino acid residues, such that X along
XX in combination with C or B does not provide a dibasic amino acid pair
XX and B is a polypeptide that can be selected from AAW53125 to AAW53126.
XX The secretory leader is a S. cerevisiae alpha-factor secretory leader. A
XX cultured yeast cell containing such an expression vector can be used to
XX produce the TPO polypeptide. The TPO polypeptide can be used in a method
XX for increasing platelet number in a mammal. It can be used to increase
XX proliferation of bone marrow cells for treatment of cytopenia, including
XX those induced by aplastic anaemia, myelodysplastic syndromes,
XX chemotherapy or congenital cytopenias. It can also be used to treat
XX thrombocytopenia, haematologic disorders, such as leukaemia and lymphoma
XX or metastatic cancers involving bone marrow.
XX
XX Sequence 1062 BP; 223 A; 375 C; 236 G; 228 T; 0 other;

Query Match 48.3%; Score 523.2; DB 19; Length 1062;
Best Local Similarity 99.4%; Pred. No. 9.8e-116;
Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATGGAGCTGACTGAATTCCTCCTCGTGTGCTATGCTTCCTCACTGCAAGCTCAACGCTG 60
DB 1 atggagctgactgaattgctcctcctcggtgctatgcttccctcaactgcaagctcaacgctg 60
QY 61 TCCAGCCCGGCTCCTCGTGTGACCTCCGAGTCCCTCAGTAACTGCTTCGTGACTCC 120
DB 61 tccagcccggtcctcctcggtgacccctcgagctccgaggttcacactgcttcgactcc 120
QY 121 CATGCTTTCACAGACTGAGCAGTGCAGAGTTCACCTTTGCCCTACACCTGTC 180
DB 121 catgcttctcacagactgagcagtgccagaggttcacacttgcctacactgtc 180
QY 181 CTGCTGCTGCTGTGAGCTTTAGCTTGGAGATGGAATGGAATGGAATGGAATGGAATGGA 240
DB 181 ctgctgctgctgtggacttttagcttggagatggagatggagatggagatggagatggag 240
QY 241 GCACAGGACATTCCTGGAGCAGTGAACCTTCTGCTGGAGGAGTGAATGGAATGGAATGGA 300
DB 241 gcacaggacattctggagcagtgacccttctgctggagagtgatggagcagcgaggga 300
QY 301 CAATGGGACCACTTGCTCTCATCCCTCTGCTGGGAGCTTTCTGGACAGGTCGCTGTC 360
DB 301 caactgggaccacttgctctcatccctctgctgggagctttctggacaggtcgtctc 360

Db 301 caactggagaccacttgctctctatccctctctggtggcagctttttgtggacaggtctcctc 360
 QY 361 CTCCTTGGGGCCCTGCAGAGCTCTCTTGAACCCAGCTTCTCCACAGGCGAGGACACACA 420
 Db 361 ctccttggggccctgcagagcctcttgaaccagcttctctccacagggcagaccaca 420
 QY 421 GCTCACAGGATCCATGCGCATCTCTGAGCTTCCACACCTGCTCCGAGGAAGGTG 480
 Db 421 gctcacaaggtatccaatgcatctctgagcttccacacactgctccgaggaaggtg 480
 QY 481 CTTTTCTCTGATGCTGTGAGGAGGCTCACCTCTGCTGCTCAGGGGTGGC 528
 Db 481 cgttctctgatgcttgttaggagggccacacctctgctcagggcggcc 528

RESULT 12

AAA51991
 ID AAA51991 standard; cDNA; 1062 BP.

AC AAA51991;

DT 04-DEC-2000 (first entry)

DE Human thrombopoietin cDNA.

XX Thrombopoietin; TPO; erythropoietin; EPO; erythropoiesis; stimulate;
 KW anaemia; radiation; chemotherapy; red blood cell production;
 KW haematopoietic cell; cytopenia; ex vivo expansion; anti-anaemic;
 KW immunomodulator; ds.
 XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..1062

FT /*tag= a

FT /product= Thrombopoietin

XX US099830-A.

PD 08-AUG-2000.

XX 12-SEP-1997; 97US-0999927.

PR 09-AUG-1994; 94US-0288417.

PR 07-NOV-1994; 94US-0335566.

XX (UNIW) UNIV WASHINGTON.

PA (ZYMO) ZYMOGENETICS INC.

XX Kaushansky K;

XX WPI; 2000-557618/51.

DR P-PSDB; AAY97136.

XX Stimulating erythropoiesis e.g. for treating anemias or cytopenia, or
 PT for ex vivo expansion of marrow or peripheral blood progenitor cell,
 PT comprises administering thrombopoietin alone or in combination with
 PT erythropoietin

XX Disclosure; Column 19-22; 23pp; English.

XX Erythropoiesis may be stimulated in anaemic mammals by administration of
 CC thrombopoietin (TPO), and optionally erythropoietin (EPO). The anaemic
 CC mammal to be treated will have shown a haematocrit level less than 33
 CC percent of total blood volume, a reticulocyte count of less than
 CC 0.8 percent and may have been treated with radiation or chemotherapy. The
 CC improved red blood cell level is due to an increase in red blood cell
 CC production rather than a decrease in red blood cell destruction. TPO
 CC proteins may be used therapeutically to increase proliferation of
 CC haematopoietic cells in the bone marrow, such as in the treatment of
 CC cytopenia or anaemia, especially associated with bone marrow failure,
 CC disorders associated with low red blood cell production, particularly
 CC when accompanied by thrombocytopenia and for increasing the level of

CC circulating erythrocytes and erythrocyte precursor cells. TPO and EPO may
 CC be used for ex vivo expansion of marrow or peripheral blood progenitor
 XX cells.
 SQ Sequence 1062 BP; 223 A; 375 C; 236 G; 228 T; 0 other;

Query Match 48.3%; Score 523.2; DB 21; Length 1062;

Best Local Similarity 99.4%; Pred. No. 9.8e-116;

Matches 525; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATTGCTTCCTCGTGGTCATGCTTCTTAAGTCAAGGCTTAACGCTG 60

Db 1 atggagctgactgaattgctctcctcggtgcatgctctctcttaactgcaaggtcaacgtg 60

QY 61 TCCAGCCCGGCTCCTCTGCTTGTGACCTCCGAGTCTCTAGTAAACGCTTCGTGACTCC 120

Db 61 tccagcccggtcctcctgctgtgacctccgagtcctcagtaactgcttcgtgactcc 120

QY 121 CATGCTCTTTCACAGCAGACTGAGCCAGTCCCGAGAGGTTTCCACCTTTGCTTACACCTGTC 180

Db 121 catgctctttcacagcagactgagccagtgcccagaggttcaccccttgcctacacctgctc 180

QY 181 CTGCTGCTGCTGTGACTTTAGCTTTGGGAGATGAAACCCAGATGGAGGAGACCAAG 240

Db 181 ctgctgctgctgtgacttttagcttgggagaatgaaaaaccagatggaggagaccaag 240

QY 241 GCACAGGACATCTGGGAGCAGTGACCTTCTCTGTGGAGGAGTGATGGCAGCACGGGGA 300

Db 241 gcacaggacattctgggagcagtgaccttctctgagggagtgatggcagcacgggga 300

QY 301 CAACTGGGACCCACTTTCCTCTCATCCCTCTGGGAGAGCTTTCCTGACAGGTCGGTCTC 360

Db 301 caactgggacccacttgcctctcatcctcctctgggcagcttcttgagacaggttcgtctc 360

QY 361 CTCCTTGGGGCCCTGCAGAGCCCTCTTGGNAACCCAGCTTCTCCACAGGCGAGGACCA 420

Db 361 ctcccttggggccctgcagagcctccttggaacccagcttctccacagggcagggaccaca 420

QY 421 GCTCACAGGATCCCAATGCCATCTTCTCGAGCTTCCACACCTGTCTCCGAGGAAAGGTG 480

Db 421 gctcacaggatcccaatgccatcttctcagcttccacacacctgctccgaggaaggtg 480

QY 481 CGTTTCTCTGATGCTTGTAGAGGTCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 528

Db 481 cgttctctgatgcttgttaggaggtccacctctgctcagggcggcc 528

RESULT 13

AAT47958

ID AAT47958 standard; cDNA to mRNA; 1086 BP.

XX AC AAT47958;

XX DT 03-APR-1997 (first entry)

XX DE Human thrombopoietin (1-332) coding sequence.

XX KW TPO; recombinant; thrombopoietin; storage; reduced adsorption;

XX KW container wall; additive; ds.

XX OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 25..1086

FT /*tag= a

FT /note= "immature protein"

FT sig_peptide 25..87

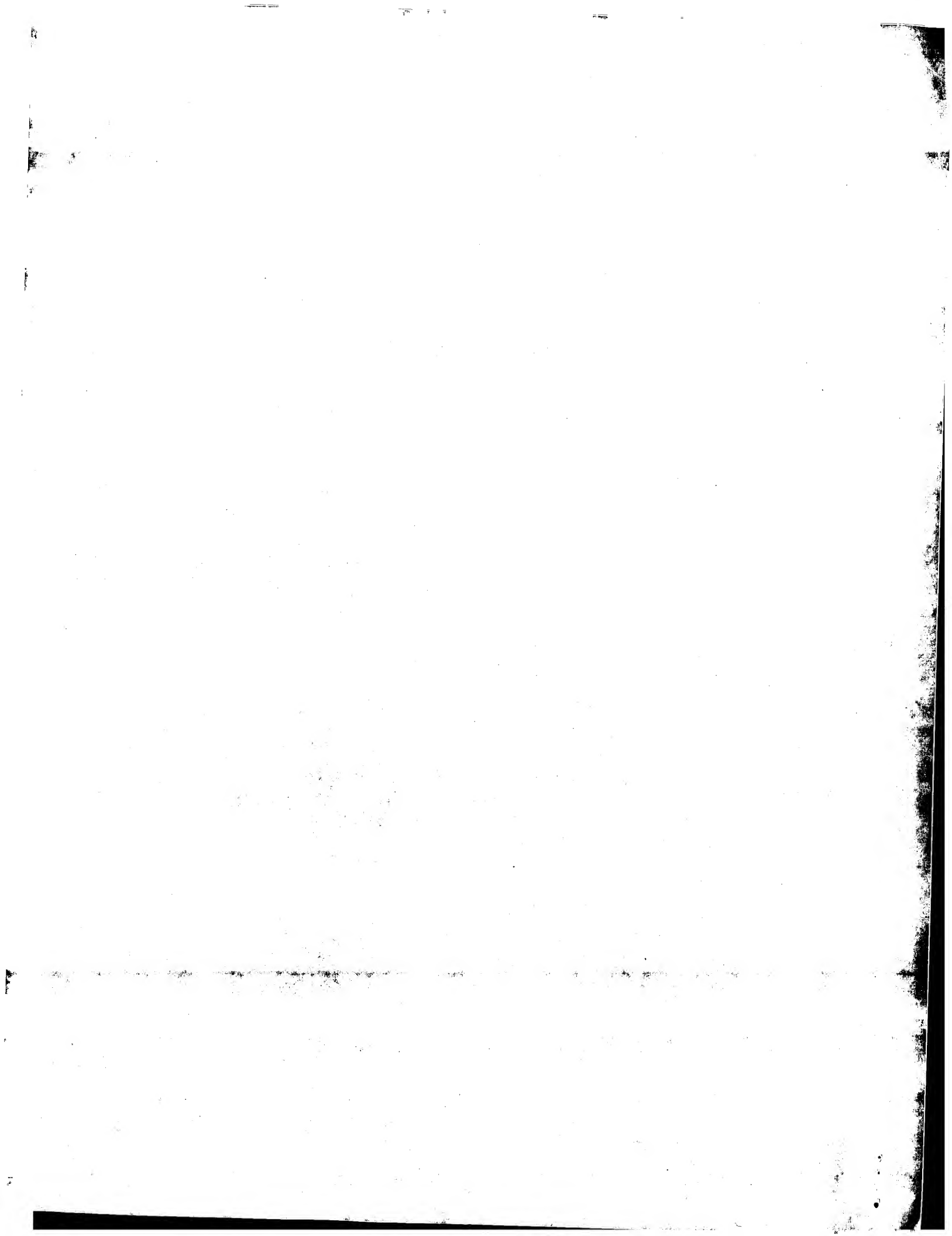
FT /*tag= b

FT mat_peptide 88..1083

FT /*tag= c

FT /note= "mature protein residues 1-332"

XX



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2002, 09:47:00 ; Search time 1909.33 seconds

(without alignments)
6095.163 Million cell updates/sec

Title: US-09-680-514-6

Perfect score: 1083

Sequence: 1 ATGGAGCTGACTGAATTGCT.....TACGCCACCTGCCAGGCC 1083

Scoring table: IDENTITY_NUC

Gap 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estro:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: gb_est1:*
11: gb_est2:*
12: gb_hic:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pin:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	320.4	29.6	973	11	BI411128
2	312.8	28.9	598	11	BG548320
3	301.6	27.8	900	11	BI328788
4	300.6	27.8	507	10	BE485194
5	300.6	27.8	633	10	AI314551
6	299.6	27.7	609	10	AA538257
7	285.2	26.3	878	11	BF789308
8	227.8	21.0	464	10	BE754587
9	135.8	12.5	327	11	BF848766
10	135.2	12.5	449	13	AZ025993
11	109.2	10.1	419	13	AZ352219
12	86.4	8.0	708	13	AZ792988
13					2M0046105

13	77.8	7.2	752	13	AQ888740
14	65.6	6.1	589	10	BE377769
15	55.6	5.1	925	13	CNS0091P
16	50.2	4.6	513	10	BE480590
17	49	4.5	925	13	CNS0091P
18	48.8	4.5	617	10	AI981598
19	43.2	4.0	461	10	BE598945
20	43.2	4.0	559	11	BG556176
21	43.2	4.0	976	11	BG823538
22	42.6	3.9	452	10	AA155632
23	42.6	3.9	899	13	CNS015Y1
24	42.4	3.9	1203	13	CNS015Y4
25	42.2	3.9	598	10	AA115932
26	41.4	3.8	421	10	AA058743
27	41.4	3.8	478	11	BI345642
28	40.8	3.8	545	11	BF868325
29	40.6	3.7	447	11	BG604563
30	40.6	3.7	845	10	AL572931
31	40.6	3.7	932	13	CNS0072Q
32	40.6	3.7	1003	11	BI410408
33	40.6	3.7	1101	13	CNS05709
34	40.4	3.7	458	10	AW273202
35	40.4	3.7	485	10	BE463718
36	40.2	3.7	864	13	CNS022LE
37	39.8	3.7	387	10	AW193322
38	39.8	3.7	427	10	BE048584
39	39.8	3.7	432	10	AW175624
40	39.8	3.7	441	10	AI818468
41	39.8	3.7	444	10	AW072844
42	39.8	3.7	450	10	AW273147
43	39.8	3.7	453	10	AW337946
44	39.8	3.7	492	13	AQ686347
45	39.8	3.7	505	10	AI952164

ALIGNMENTS

RESULT 1
BI411128
LOCUS 602962472F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5118185 5',
DEFINITION mRNA sequence.
ACCESSION BI411128.1 GI:15172051
VERSION BI411128
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 973)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLAM11287 row: O column: 18
High quality sequence start: 32
High quality sequence stop: 912.
FEATURES
Location/Qualifiers
1..973
/organism="Mus musculus"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:5118185"

```

/clone_lib="NCI_CGAP_Lu33"
/tissue_type="pooled lung tumors"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st
strand cDNA was prepared from mRNA obtained from pooled
lung tumors with a Not I - oligo(dT) primer [5'
TGTACCAATCTGAAGTGGAGCGCGCTCTGTCTTTTCTTTTCTTTT
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT      192 a      295 c      254 g      232 t
ORIGIN
Query Match      29.6%; Score 320.4; DB 11; Length 973;
Best Local Similarity 77.1%; Pred. No. 1.1e-62;
Matches 390; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
QY 578 CCTCGAGCTTACACAGAGCTTCTTTTAAAGCTTAGACAACTGAGGAAGATCCAGG 637
DB 91 CCCTGCTCTGCCCCGAAGCTTCTGCTTAAGTCCCTGGAGCAAGTGAGGAAGATCCAGG 150
QY 638 GCGATGGCGGAGCGCTCCAGGAAAGCTGTGTGCCACCTACAAAGCTGTGCCACCCCGAGG 697
DB 151 CCAGCGGCTGGTCTGCTGAGCAGTGTGTGCCACCTACAAAGCTGTGTGCCACCCGAGG 210
QY 698 AGCTGGTGTGCTGGACACTCTCTGGGCATCCCTGGGCTCCCTGGAGCAGCTGCCCA 757
DB 211 AGCTGGTGTGCTGGGCACTCTCTGGGATCCGGAAGCTTCCCTGAGTGGCTGCTTA 270
QY 758 GCCAGGCGCTGACGCTGGAGGCTGTGTGAGCAACTCCATAGCGGCTTTCCCTTACC 817
DB 271 GCCAGGCGCTGACGAGCAGACAGTGTGCTTAAGTCCCTGGAGCAAGTGAGGCTTGCCTTACC 330
QY 818 AGGCGCTCTGACGCGCTGGAAGGATCTCCCGGAGTGTGGTCCACCTTGGACACAC 877
DB 331 AAGTCTCTGCTGAGGCTTATCGGATATTTCCCTGCCCTGGCCCCACCTTGGACTTGC 390
QY 878 TGCAGCTGGAGCTGGCGAGCTTTGCCACCACTATCTGGCAGCAGATGGAAGAACTGGAA 937
DB 391 TTACGCTGGATGTGCCAACTTTGCCACCACTATCTGGCAGCAGATGGAAGAACTAGGGG 450
QY 938 TGGCGCTGCTGACGCGGCTGAGGCTGCGATGCGGCTTGCCTTGCCTTCCAGC 997
DB 451 TGGCGCTGCTGACGCGGCTGAGGCTGCGATGCGGCTTGCCTTGCCTTCCAGC 510
QY 998 GCGGCGGAGGCGCTCTAGTTGCTTCCCTGCTGAGAGCTTCCCTGGAGGTGCTGCTACC 1057
DB 511 GCGGCGGAGGCGCTCTAGTTGCTTCCCTGCTGAGAGCTTCCCTGGAGGTGCTGCTACC 570
QY 1058 GCGTCTTACGCCACCTTGGCCGAGCC 1083
DB 571 TTGCTCTGCACCACTTGGCTAGACC 596

RESULT 2
BG548320      598 bp      mRNA      EST      04-APR-2001
LOCUS      602575289F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4703159 5',
DEFINITION      mRNA sequence.
ACCESSION      BG548320
VERSION      BG548320.1 GI:13546985
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 598)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)

```

```

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLC1541 row: b column: 24
High quality sequence stop: 597.
Location/Qualifiers
1..598
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_77"
/lab_host="DH10B (T1 phage-resistant)"
/notes="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggccattatgcc); Site_2: SfiI (ggccattatgcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGCCATTATGGC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGGCGGCGGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.9
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."
BASE COUNT      113 a      202 c      163 g      120 t
ORIGIN
Query Match      28.9%; Score 312.8; DB 11; Length 598;
Best Local Similarity 99.4%; Pred. No. 5.1e-61;
Matches 314; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 768 GCAGCTGGCAGGCTGTGTGAGCCAACTCCATAGCGGCTTTCTCTACAGGGGCTCCT 827
DB 185 GAAGCTGGCAGGCTGTGTGAGCCAACTCCATAGCGGCTTTCTCTACAGGGGCTCCT 244
QY 828 GCAGGCGCTTGAAGGATCTCCCGGAGTGGGTGCCACACTTGGACACACTGCAGCTGGA 887
DB 245 GCAGGCGCTTGAAGGATCTCCCGGAGTGGGTGCCACACTTGGACACACTGCAGCTGGA 304
QY 888 CGTGGCGGCTTGGCCACCACTCTGGCAGCAGATGGAAGAACTGGGAATGGCCCTGC 947
DB 305 CGTGGCGGCTTGGCCACCACTCTGGCAGCAGATGGAAGAACTGGGAATGGCCCTGC 364
QY 948 CTGTCAGCCACCCAGGCTGCCATCGCGCTTGGCTTCTGCTTCCAGCGCGGCGAGG 1007
DB 365 CTGTCAGCCACCCAGGCTGCCATCGCGCTTGGCTTCTGCTTCCAGCGCGGCGAGG 424
QY 1008 AGGGTCTCTAGTTGCTTCCCTCCATCTGCAGAGCTTCTCTGGAGGTGCTGCTTCTACG 1067
DB 425 AGGGTCTCTAGTTGCTTCCCTCCATCTGCAGAGCTTCTCTGGAGGTGCTGCTTCTACG 484
QY 1068 CCACCTTGGCCGAGCC 1083
DB 485 CCACCTTGGCCGAGCC 500

RESULT 3
BI328788
LOCUS      602984776F1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:5137578 5',
DEFINITION      mRNA sequence.
ACCESSION      BI328788
VERSION      BI328788.1 GI:15013445
KEYWORDS      EST.
SOURCE      house mouse.
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Accession	Sequence	Position
Db	CGCCATCCGACCTTCACTTACGCCCTTCCACGCAGACGAGAGGGTCTGGTTGCTTC	463
QY	CCATCTCGAGAGTTCCTGGAGGTGTGTCATCCGCGTTCTACGC	1068
Db	CCAGCTGCATCGTTTCTTGGAGCTGGCATACCGTGGCGCTGCGC	506


```
Query Match      27.7%; Score 299.6; DB 10; Length 609;
Best Local Similarity 81.6%; Pred. No. 5.1e-58;
Matches 364; Conservative 0; Mismatches 69; Indels 13; Gaps 1;

QY 1 ATGGAGCTGACTGAATTCCTCGTGGTGCATGCTTCTCTAACTGCAAGGCTAAACGCTG 60
D 177 ATGGAGCTGACTGATTTGCTCTCTGGGCGCATGCTTCTTGCAGTGGCAAGACTAACTCTG 236
QY 61 TCAGAGCCGCTCTCTGCTGCTGAGCTCCGAGTCTCAGTAACTGCTTGTGACTCC 120
D 237 TCAGAGCCGCTGAGTCTCTGCTGCTGAGCTCCGAGTCTCAGTAACTGCTTGTGACTCC 296
QY 121 CATGTCTCTTACAGCAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 180
D 297 CACTCTCTTACAGCAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 356
QY 181 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
D 357 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 416
QY 241 GCACAGGACATTCAGGAGCAGTGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 300
D 417 GCACAGGACATTCAGGAGCAGTGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 476
QY 301 CAATGGGACCCACTTGCCTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
D 477 CAGTTGGAACCTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 536
QY 361 CTCTTGGGCGCTTGCAGAGCTCTCTTGGAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
D 537 CTCTTGGGCGCTTGCAGAGCTCTCTTGGAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 583
QY 421 GCTCAGAGATCCCAATGCTCTT 446
D 584 GCTCAGAGATCCCAATGCTCTT 609

RESULT 7
BF789308 878 bp mRNA EST 12-JAN-2001
LOCUS 602104729F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:422927
DEFINITION 5', mRNA sequence.
ACCESSION BF789308
VERSION BF789308.1 GI:12094344
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Mus.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing: Incyte Genomics, Inc.
CDNA Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9810 row: 1 column: 08
High quality sequence stop: 628.
Location/Qualifiers
1. 878
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:422927"
/clone_lib="NCI_CGAP_Kid14"
```

```
/lab_host="DH10B (T1 phage-resistant)"
/notes="Organ: kidney; Vector: pCMV-SPORT6; Site: 1; NotI;
Site: 2; Sali; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library. 1"
BASE COUNT 203 a 281 c 224 g 169 t 1 others
ORIGIN

Query Match      26.3%; Score 285.2; DB 11; Length 878;
Best Local Similarity 84.7%; Pred. No. 1e-54;
Matches 320; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATTCCTCGTGGTGCATGCTTCTCTAACTGCAAGGCTAAACGCTG 60
D 238 ATGGAGCTGACTGATTTGCTCTCTGGGCGCATGCTTCTTGCAGTGGCAAGACTAACTCTG 297
QY 61 TCAGAGCCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
D 298 TCAGAGCCGCTGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 357
QY 121 CATGTCTCTTACAGCAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 180
D 358 CACTCTCTTACAGCAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 417
QY 181 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
D 418 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 477
QY 241 GCACAGGACATTCAGGAGCAGTGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 300
D 478 GCACAGGACATTCAGGAGCAGTGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 537
QY 301 CAATGGGACCCACTTGCCTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
D 538 CAGTTGGAACCTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 597
QY 361 CTCTTGGGCGCTTGCAG 378
D 598 CTCTTGGGCGCTTGCAG 615

RESULT 8
BF754587 464 bp mRNA EST 25-APR-2001
LOCUS 208137 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION BF754587
ACCESSION BF754587
VERSION BF754587.1 GI:10168579
KEYWORDS EST.
SOURCE cow.
ORGANISM Bos taurus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
TITLE 1 (bases 1 to 464)
JOURNAL Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,
MEDLINE Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett
COMMENT , G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G.,
Perle, G., Holt, I., Karameycheva, S., Liang, F., Quackenbush, J. and
Keele, J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
21180013
CONTACT: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
```

PCR Primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCCTCAGTCAGCAGC
 Plate: 54 row: J column: 23
 Seq primer: ATTAGGTGACACATATAG.

FEATURES

Location/Qualifiers
 1. .464
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /clone_lib="MARC 2BOV"
 /tissue_type="pooled"
 /lab_host="DH10B"

/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
 Library made from pooled tissue from testis, thymus,
 semitendinosus muscle, longissimus muscle, pancreas,
 adrenal, and endometrium." 87 t

BASE COUNT 97 a 152 c 128 g 87 t
 ORIGIN

Query Match 21.0%; Score 227.8; DB 10; Length 464;
 Best Local Similarity 87.2%; Pred. No. 9.7e-42;
 Matches 266; Conservative 0; Mismatches 27; Indels 12; Gaps 1;

Qy 203 GCTTGGAGAAATGGAACCCAGATGGAGAGACCAAGCAGACGACATCTCTGGAGCAG 262
 Db 1 GCTTGGAGAAATGGAACCCAGACGAGCAGACCAAGCAGACGACGCTCTGGGAACCA 60
 Qy 263 TGACCTCTCTGCTGGAGGAGTATGGCAGCAGCGGGGACAACTGGGACCCACTTGCCTCT 322
 Db 61 CGACCTCTCTGCTGGAGGAGTATGACAGCGGGGAGCAGCTGGGCCCCACTTGCCTCT 120
 Qy 323 CATCCCTCTCTGGGCGAGCTTCTTGGACAGGTCCTCTCTCTTGGGGCCCTGCAGAGCC 382
 Db 121 CATCCCTCTCTGGGCGAGCTTCTTGGCAGTCGCGCTCTCTTGGGGCCCTGCAGGGCC 180
 Qy 383 TCCTTGAACCCAGCTTCTCTCAGGCGAGGACCAACAGCTCACAAGGATGCCAATGCCA 442
 Db 181 TCTTAGAACCC-----CAGGCGAGGACCAACAGCTCACAAGGATGCCAAGTGCCA 228
 Qy 443 TCTTCTGAGCTTCCCAACACCTGCTCCGAGAAAGTGCGTTTCTCTGATGCTTGTAGGAG 502
 Db 229 TCTTCTGAGCTTCCCAACAGCTGCTCCGAGAAAGTGCGCTTCTCTGCTTGTAGTG 288
 Qy 503 GGTC 507
 Db 289 GGGCC 293

RESULT 9
 BF848766 327 bp mRNA EST 16-JAN-2001
 LOCUS QVO-EN0102-081100-458-g06 EN0102 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BF848766
 ACCESSION BF848766
 VERSION BF848766.1 GI:12235903
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 327)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 JOURNAL
 MEDLINE
 COMMENT 20202663
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?ti=QV0&t2=QV0-EN0102-
 081100-458-g06&t3=2000-11-08&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 327.

FEATURES

Location/Qualifiers
 1. .327
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="EN0102"
 /dev_stage="Adult"

/note="Organ: lung_normal; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions." 76 t

BASE COUNT 61 a 87 c 103 g 76 t
 ORIGIN

Query Match 12.5%; Score 135.8; DB 11; Length 327;
 Best Local Similarity 89.6%; Pred. No. 7.1e-21;
 Matches 146; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 762 GGCCTGCGAGCTGGAGGCTGCTTGGACCAATCCATAGCGGCTTTCTCTACACAGG 821
 Db 139 GGCCTATACAGAGCCCGGCTGTTGAGCAATCCATAGCGGCTTTCTCTACACAGG 198
 Qy 822 GTCCTGCGAGCCCTGGAAGGATCTCCCGGAGTTGGTCCACCTTGGACACACTGCA 881
 Db 199 GCTCTGCGAGCCCTGGAAGGATCTCCCGGAGTTGGTCCACCTTGGACACACTGCA 258
 Qy 882 GCTGACGCTCGCGACTTTCACACCACTCTGCGAGCAGATG 924
 Db 259 GCTGACGCTCGCGACTTTCACACCACTCTGCGAGCAGATG 301

RESULT 10
 AZ025993 449 bp DNA GSS 25-FEB-2000
 LOCUS RPCI-23-343116.TJ RPCI-23 Mus musculus genomic clone RPCI-23-343116
 DEFINITION , DNA sequence.
 ACCESSION AZ025993
 VERSION AZ025993.1 GI:7101377
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 449)
 Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akinret,
 B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P.
 and Fraser, C.M.
 Mouse BAC End Sequences from Library RPCI-23
 Unpublished (1999)
 Other GSSs: RPCI-23-343116.TV
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC

library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>) or from Resea ch Genetics (http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html). BAC end page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html

Plate: 343 row: I column: 16
Seq primer: SP6
Class: BAC ends.

FEATURES

Location/Qualifiers
1..449
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-343116"
/clone_lib="RPCI-23"
/sex="Female"
/lab_host="DH10B"

/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site:1: EcoRI; Site:2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 90 a 121 c 108 g 129 t 1 others
ORIGIN

Query Match 12.5%; Score 135.2; DB 13; Length 449;
Best Local Similarity 84.4%; Pred. No. 1e-20;
Matches 152; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 228 GGAGGAGACCAAGCAGACAGATCTGGGAGCAGTACCTCTCTGCTGGAGGAGTGAT 287
Db 126 GGAACAGCAGACAGCAGACATCTAGGGGAGTCTTTCTACTGAGGAGTGAT 185

QY 288 GGCAGCAGGGGACAACTGGGAGCCACTTGCCTCTCATCCCTCTCTGGGCGAGCTTTCTGG 347
Db 186 GGCAGCAGGAGCAGAGTGGACCCCTCTCTCATCCCTCTCTGGGAGCTTTCTGG 245

QY 348 ACAGGTCCGTCCTCTCTGGGCGCTTCGAGAGCCTCTTGGACCCAGCTTCCTCCACA 407
Db 246 GCAGGTTCGCTCTCTGGGCGCTTCGAGAGCCTCTTGGACCCAGCTTCCTCCACA 305

RESULT 11

AZ352219/c
LOCUS 419 bp DNA GSS 29-SEP-2000
DEFINITION 1M0090P05R Mouse 10kb plasmid UUGCLM library Mus musculus genomic clone UUGCLM0090P05 R, DNA sequence.

ACCESSION AZ352219
VERSION AZ352219.1 GI:10431456
KEYWORDS GSS.

SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 419)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00
Plate: 0090 row: P column: 05
Seq primer: CACACAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 419.
Location/Qualifiers
1..419

FEATURES

source
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGCLM0090P05"
/clone_lib="Mouse 10kb plasmid UUGCLM library"
/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gil4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 100 a 98 c 124 g 97 t
ORIGIN

Query Match 10.1%; Score 109.2; DB 13; Length 419;
Best Local Similarity 79.6%; Pred. No. 8.3e-15;
Matches 129; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 360 CCTCTTGGGCGCTTCGAGAGCCTCTTGGAGCCAGCTTCTCCACAGGCGGACCCAC 419
Db 292 CCTACTCACTGGCTCAGGCCCATCTCTGCCCTCAGCTTCTCTACAGGCGGACCCAC 233

QY 420 AGCTCACAAGGATCCCAATGCCATCTCTCTGAGCTTCCACACACCTGCTCCGAGAAAGT 479
Db 232 AGCTCACAAGGATCCCAATGCCATCTCTCTGAGCTTCCACACACCTGCTCCGAGAAAGT 173

QY 480 GCGTTTCTCATGCTTCTAGGAGGTCACCCCTCTGCGTCAG 521
Db 172 GCGCTTCTCTGCTGCTAGAGGTCACCCCTCTGCTGTCAG 131

RESULT 12

AZ792988
LOCUS 708 bp DNA GSS 16-FEB-2001
DEFINITION 2M0046105F Mouse 10kb plasmid UUGCLM library Mus musculus genomic clone UUGCLM0046105 F, DNA sequence.

ACCESSION AZ792988
VERSION AZ792988.1 GI:12937676
KEYWORDS GSS.

SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 708)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM8767 row: g column: 06
High quality sequence stop: 587.
Location/Qualifiers
1. .589
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_lib="NCI_CGAP_Maml"
/tissue_type="tumor, biopsy sample"
/dev_stage="3 months, virgin"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

FEATURES

source

150 a 152 c 134 g 153 t

BASE COUNT
ORIGIN

120 a 61 c 172 t 511 others

Query Match 6.1%; Score 65.6; DB 10; Length 589;
Best Local Similarity 76.9%; Pred. No. 7.2e-05;
Matches 80; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
Qy 1 ATGAGCTGACTGATTCCTCGTGTGATGCTTCCTCACTGCAAGCTCAAGCTG 60
|||||
Db 484 ATGAGCTGACTGATTCCTCGTGTGATGCTTCCTCACTGCAAGCTCAAGCTG 543
|||||
Qy 61 TCCAGCCCGGCTCTCTGCTGACCTCGAGCTCGAGCTCAGTAA 104
|||||
Db 544 TCCAGCCCGGCTCTCTGCTGACCTCGAGCTCGAGCTCAGTAA 587
|||||

RESULT 15

CNS0091P 925 bp DNA GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL053013.1 GI:4934461
VERSION GSS.
KEYWORDS fruit fly,
SOURCE Drosophila melanogaster
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 925)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osogawa and
Aron Mammoser in Pletier de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's

REFERENCE

AUTHORS
TITLE
JOURNAL

COMMENT

P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

1. .925

/organism="Drosophila melanogaster"

/db_xref="taxon:7227"

/clone_lib="RPCI-98"

/clone="BACR19D16"

/note="end : TET3"

120 a 61 c 172 t 511 others

Query Match 5.1%; Score 55.6; DB 13; Length 925;
Best Local Similarity 16.1%; Pred. No. 0.015;
Matches 74; Conservative 175; Mismatches 208; Indels 3; Gaps 1;
Qy 556 TCTAGAGCACCAATATCGCGCCTCGAGTCTACACAGAGCTTCCTTTTAAAGCTTA 615
|||||
Db 457 TATTAATAATAATTTTNTNNNNNNNNNAATANANNNTTATTATTANNNNANNA 516
|||||
Qy 616 GAGCAAGTGAAGAAATCCAGGGCGATGGCGAGCTCCAGGAGAACTGTGTCCACC 675
|||||
Db 517 NANANANNNNAGCSMSMCKGKCGSTTBGTTTTTTTTSSGGYKGCSSGGBSCSCS 576
|||||
Qy 676 TACAAGTGTGCCACCGGAGAGCTGTGTCTGTCGACACTCTCTGGGATCCCTGG 735
|||||
Db 577 SCSCSSSSCCBCCCSCSSCYCCSSBSBSKCSSTBSBSCSSKSVGCTSCSSSS 636
|||||
Qy 736 GCTCCCTGACAGCTGCCCGCAGCCCTGCGAGCTGCGAGCTGCTTTCAGCAACTC 795
|||||
Db 637 CSSSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSS 696
|||||
Qy 796 CATAGCGGCTTTTCTCTTACAGGGCTCTCTGAGGCTTCTGAGAGGATCTCCCGAG 855
|||||
Db 697 SSSSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSS 756
|||||
Qy 856 TTGGGTCCCACTTGGACACACTGCA--GCTGGACGTCGCCGACTTGCACACCATC 912
|||||
Db 757 TCCTCCCTCCSYSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSS 816
|||||
Qy 913 TGGCAGCAGATGGAAGAACTGGGAATGCCCTGCTGAGCCACCCAGGGTGCCTATG 972
|||||
Db 817 TYBMBCTTSCGSSSSSSSGKGVTKCGCGGSSSTNGMBGTSSACSSSSSSSSSVS 876
|||||
Qy 973 CCGGCTTCGCTCTGCTTTCAGCGCGGCGGAGAGGGG 1012
|||||
Db 877 SSKKSSASSSSSSSGSGSVSSSSASAKSSSSGVS SSGS 916
|||||

Search completed: April 5, 2002, 09:47:02
Job time: 3326 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2002, 10:29:34 ; Search time 2520.15 seconds
(without alignments)
7167.991 Million cell updates/sec

Title: US-09-680-514-8

Perfect score: 1095

Sequence: 1 ATGGAGCTGACTGAATTGCT.....TAGGCCACCTTGCCAGCCC 1095

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 824859755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_em.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_om.*

20: em_or.*

21: em_ov.*

22: em_pat.*

23: em_ph.*

24: em_pl.*

25: em_ro.*

26: em_sts.*

27: em_sy.*

28: em_un.*

29: em_vi.*

30: em_htgo_hum.*

31: em_htgo_inv.*

32: em_htgo_rod.*

33: em_htg_hum.*

34: em_htg_inv.*

35: em_htg_rod.*

36: em_htg_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

8

Result No.	Score	Match	Length	DB	ID	Description
1	520.4	47.5	525	6	E02574	E02574 DNA encodin
2	518.8	47.4	555	22	E11965	E11965 Human cDNA
3	518.8	47.4	1059	6	AR103959	AR103959 Sequence
4	518.8	47.4	1062	6	AR087133	AR087133 Sequence
5	518.8	47.4	1062	6	AR087571	AR087571 Sequence
6	518.8	47.4	1062	6	I49760	I49760 Sequence 3
7	518.8	47.4	1062	9	H0MTHROMB	L36052 Human throm
8	518.8	47.4	1086	6	E16669	E16669 cDNA encodi
9	518.8	47.4	1267	6	E12182	E12182 Human cDNA
10	518.8	47.4	1267	6	E12214	E12214 Human cDNA
11	518.8	47.4	1342	6	AR008878	AR008878 Sequence
12	518.8	47.4	1342	6	AR023468	AR023468 Sequence
13	518.8	47.4	1342	6	AR087573	AR087573 Sequence
14	518.8	47.4	1342	6	AR087573	AR087573 Sequence
15	518.8	47.4	1691	9	D32047	D32047 Homo sapien
16	518.8	47.4	1721	6	E16668	E16668 cDNA encodi
17	518.8	47.4	1721	22	E11961	E11961 Human cDNA
18	518.8	47.4	1795	6	I33525	I33525 Sequence 1
19	518.8	47.4	1795	9	H0MMLCMLP	L33410 Human c-mpl
20	517.2	47.2	605	6	AR008879	AR008879 Sequence
21	517.2	47.2	605	6	AR087574	AR087574 Sequence
22	517.2	47.2	605	6	I85130	I85130 Sequence 3
23	515.6	47.1	1341	9	HSU11025	U11025 Human megak
24	514	46.9	525	6	E02575	E02575 DNA encodin
25	514	46.9	1062	9	HSU59493	U59493 Human throm
26	514	46.9	1062	9	HSU59494	U59494 Human throm
27	512.4	46.8	525	6	E02576	E02576 DNA encodin
28	486.8	44.5	525	6	AR024358	AR024358 Sequence
29	486.8	44.5	525	6	AR091731	AR091731 Sequence
30	486.8	44.5	525	6	E01731	E01731 cDNA encodi
31	486.8	44.5	525	6	E02573	E02573 DNA encodin
32	486.8	44.5	525	6	E07164	E07164 cDNA encodi
33	486.8	44.5	525	6	E15131	E15131 Human mRNA
34	486.8	44.5	525	6	I71150	I71150 Sequence 1
35	486.8	44.5	525	6	I83709	I83709 Sequence 1
36	486.8	44.5	525	22	E09431	E09431 cDNA encodi
37	486.8	44.5	660	9	H0MGC5F	M17706 Human granu
38	486.8	44.5	1498	9	HSGC5F1	X03655 Human mRNA
39	486.8	44.5	1521	6	E01219	E01219 cDNA encodi
40	486.8	44.5	1525	6	I05333	I05333 Sequence 11
41	486.8	44.5	1525	6	I08669	I08669 Sequence 1
42	486.8	44.5	1525	6	I09205	I09205 Sequence 5
43	486.8	44.3	522	6	E01631	E01631 DNA encodin
44	485.2	44.3	522	6	E08529	E08529 DNA encodin
45	485.2	44.3	1414	9	H0MGC5F	M13008 Human granu

ALIGNMENTS

RESULT	1					
E02574						
LOCUS	E02574	DNA	525 bp	DNA	PAT	29-SEP-1997
DEFINITION	DNA encoding human colony-stimulating factor derivative.					
ACCESSION	E02574					
VERSION	E02574.1	GI:2170804				
KEYWORDS	JP 1990227075-A/2.					
SOURCE	synthetic construct.					
ORGANISM	synthetic construct					
REFERENCE	1 (bases 1 to 525)					
AUTHORS	Sasaki, K., Nishi, T., Yasumura, S., Sato, M. and Itou, S.					
TITLE	NEW POLYPEPTIDE					
JOURNAL	Patent: JP 1990227075-A 2 10-SEP-1990;					
COMMENT	OS Artificial gene					
	OC Artificial sequence; Genes.					
	PN JP 1990227075-A/2					
	PD 10-SEP-1990					
	PF 28-SEP-1989 JP 1989253097					

```
PR 29-SEP-1988 JP 88P 245705
PI SASAKI KATSUTOSHI, NISHI TATSUYA, YASUMURA SHIGEYOSHI, PI
SATO MORIYUKI,
PI ITOU SEIGA
PC C12N9/72,C07K13/00,C07K15/14,C12N1/21,C12N5/10,C12N9/64, PC
C12N15/27,
PC C12N15/58,C12N15/70,C12N15/85,C12P21/02,C12P21/02,(C12N1/21,
C12R1:19),
PC (C12P21/02,C12R1:19),(C12P21/02,C12R1:91);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FH mat_peptide 1..522
FT /product='human colony-stimulating factor' FT
FT derivative which
FT is named 'hg-CSF[ND28]'.
FEATURES
source
Location/Qualifiers
1..525
/organism='synthetic construct'
BASE COUNT 94 a 184 c 148 g 99 t
ORIGIN
Query Match 47.5%; Score 520.4; DB 6; Length 525;
Best Local Similarity 99.8%; Pred. No. 3.3e-96;
Matches 521; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 574 GCGCCACATATCGCGCTCGAGTCTACCACAGAGCTTCTTTTAAAGCTTAGAGCAA 633
Db 1 GCACACATATCGCGCTCGAGTCTACCACAGAGCTTCTTTTAAAGCTTAGAGCAA 60
Qy 634 GTGAGGAGATCCAGGGCGATGGCGAGCGCTCCAGAGAGAGCTGTGTGCCACTACAAG 693
Db 61 GTGAGGAGATCCAGGGCGATGGCGAGCGCTCCAGAGAGAGCTGTGTGCCACTACAAG 120
Qy 694 CTGTGCCACCCCGAGGAGCTGTGTCTCGGACACTCTCTGGGCATCCCTGGGCTCCC 753
Db 121 CTGTGCCACCCCGAGGAGCTGTGTCTCGGACACTCTCTGGGCATCCCTGGGCTCCC 180
Qy 754 CTGAGCAGCTGCCCGACCCAGCGCCCTGCAGCTGGCAGCTGTGTGAGCCAACTCCATAGC 813
Db 181 CTGAGCAGCTGCCCGACCCAGCGCCCTGCAGCTGGCAGCTGTGTGAGCCAACTCCATAGC 240
Qy 814 GGCCTTTTCTCTACAGGGGCTCTGCGAGGCGCTCGAAGGGATCTCCCGCGAGTTGGGT 873
Db 241 GGCCTTTTCTCTACAGGGGCTCTGCGAGGCGCTCGAAGGGATCTCCCGCGAGTTGGGT 300
Qy 874 CCCACCTTGGACACACTGCAGCTGGAGTGGCGGACTTTGCCACCACTCTGGCAGCAG 933
Db 301 CCCACCTTGGACACACTGCAGCTGGAGTGGCGGACTTTGCCACCACTCTGGCAGCAG 360
Qy 934 ATGAAGAAGCTGGAATGGCCCTGCGCTGCAGCCACCCAGGGTGCCATGCCGCGCTTC 993
Db 361 ATGAAGAAGCTGGAATGGCCCTGCGCTGCAGCCACCCAGGGTGCCATGCCGCGCTTC 420
Qy 994 GCCTCTGCTTCCAGCGCGGGGAGGAGGGTCTAGTGGCTTCCATCTGCAGAGCTTC 1053
Db 421 GCCTCTGCTTCCAGCGCGGGGAGGAGGGTCTAGTGGCTTCCATCTGCAGAGCTTC 480
Qy 1054 CTGAGGAGTGTGTAACCGCTTCTACGCCACCTTGGCCAGGCC 1095
Db 481 CTGAGGAGTGTGTAACCGCTTCTACGCCACCTTGGCCAGGCC 522
RESULT 2
ID E11965
XX standard; RNA; HUM; 555 BP.
AC E11965;
```

```
XX E11965.1
SV
XX
XX
DT 08-OCT-1997 (Rel. 52, Created)
DT 02-SEP-2000 (Rel. 65, Last updated, version 2)
XX
XX Human cDNA encoding thrombopoietin.
KW JP 1996228781-A/6.
XX
OS Homo sapiens (human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX
XX [1]
RN 1-555
RP Miyazaki H., Kato H., Ogami K., Iwamatsu A., Akahori H., Kuroki R.,
RA Shimizu T., Muto T.;
RT "DNA CODING FOR PROTEIN HAVING TPO ACTIVITY";
RL Patent number JP1996228781-A/6, 10-SEP-1996.
RL KIRIN BREWERY CO LTD.
XX
CC OS Homo sapiens (human)
CC PN JP 1996228781-A/6
CC PD 10-SEP-1996
CC PF 14-FEB-1995 JP 1995063298
CC PR 14-FEB-1994 JP 94P 39090, 25-MAR-1994 JP 94P 79842,
CC PR 01-JUN-1994 JP 94P 155126, 15-JUN-1994 JP 94P 167328,
CC PR 17-AUG-1994 JP 94P 227159, 17-AUG-1994 JP 94P 193169,
CC PR 01-NOV-1994 JP 94P 304167, 01-DEC-1994 JP 94P 298669,
CC PR 28-DEC-1994 JP 94P 341200
CC PI MIYAZAKI HIROSHI, KATO HISASHI, OGAMI KINYA, IWAMATSU AKIHIKO,
CC PI AKAHORI HIRONORI, KUROKI RYOTA, SHIMIZU TOSHIYUKI,
CC PI MUTO TAKANORI
CC PC C12N15/09,A61K38/00,C07K14/52,C12N1/21,C12N5/10,C12P21/02,
CC PC (C12N1/21,
CC PC C12R1:19),(C12N5/10,C12R1:91),(C12P21/02,C12R1:19),(C12P21/02,
CC PC C12R1:91);
CC CC strandedness: Double;
CC CC topology: Linear;
CC FH Key Location/Qualifiers
CC FH source 1..555
CC FT /organism='Homo sapiens'
CC FT /tissue_type='liver'
CC FT CDS 1..555
CC FT /product='thrombopoietin'
CC FT sig_peptide 1..63
XX
XX Key Location/Qualifiers
FH source 1..555
FH /db_xref='taxon:9606'
FH /organism='Homo sapiens'
XX
SQ Sequence 555 BP; 106 A; 183 C; 145 G; 121 T; 0 other;
Query Match 47.4%; Score 518.8; DB 22; Length 555;
Best Local Similarity 99.6%; Pred. No. 7e-96;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGGAGCTGACTGAATTGCTCGTGGTTCATGCTTCTCTTAAGCTCAAGCTAACGCTG 60
Db 1 ATGGAGCTGACTGAATTGCTCGTGGTTCATGCTTCTCTTAAGCTCAAGCTAACGCTG 60
Qy 61 TCCAGCCCGGCTCCTCTCTGCTTGTGACCTCCGAGTCCAGTAACTGCTTCGTGACTCC 120
Db 61 TCCAGCCCGGCTCCTCTCTGCTTGTGACCTCCGAGTCCAGTAACTGCTTCGTGACTCC 120
Qy 121 CATGTCTTTCACAGCAGACTGAGCCAGTCCAGAGGTTCCACCTTTGCCCTACACCTGTC 180
Db 121 CATGTCTTTCACAGCAGACTGAGCCAGTCCAGAGGTTCCACCTTTGCCCTACACCTGTC 180
```

Qy	301	CAACTGGGACCCACTTGCCTCTCATATCCCTCTCTGGGCGAGCTTCTTGACAGGTCCTGCTC	360
Db	301	CAACTGGGACCCACTTGCCTCTCATATCCCTCTCTGGGCGAGCTTCTTGACAGGTCCTGCTC	360
Qy	361	CTCCTTGGGGCCCTGCAGAGCCTCTTGGRACCCAGCTTCTCCACAGGCGAGGACCACA	420
Db	361	CTCCTTGGGGCCCTGCAGAGCCTCTTGGRACCCAGCTTCTCCACAGGCGAGGACCACA	420
Qy	421	GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCCAACACCTGCTCCGAGGAAAGGTG	480
Db	421	GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCCAACACCTGCTCCGAGGAAAGGTG	480
Qy	481	CGTTTCCTGATGCTTGAGGAGGTCACCCCTCTGCGTAGCG	522
Db	481	CGTTTCCTGATGCTTGAGGAGGTCACCCCTCTGCGTAGCG	522

RESULT	4
AR087133	
LOCUS	AR087133 1062 bp DNA
DEFINITION	Sequence 3 from patent US 5986049.
ACCESSION	AR087133
VERSION	AR087133.1 GI:10013899
KEYWORDS	.
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 1062)
TITLE	Forstrom,J.W., Lofton-Day,C.E. and Lok,S.
JOURNAL	Purified thrombopoietin and method of making it
FEATURES	Patent: US 5986049-A 3 16-NOV-1999;
source	Location/Qualifiers
	1..1062

Query Match	47.4%	Score 518.8;	DB 6;	Length 1062;
Best Local Similarity	99.6%;	Pred. No. 5.8e-96;		
Matches 520;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0
Qy	1	ATGAGCTGACTGAATTGCTCTCTCGTGTGTCATGCTTCTCTCAACTGCAAGGCTAAACGCTG	60	
Db	1	ATGAGCTGACTGAATTGCTCTCTCGTGTGTCATGCTTCTCTCAACTGCAAGGCTAAACGCTG	60	
Qy	61	TCCAGCCCGGCTCCTCCTTGCTGTGACCTCCGAGTCCCTCAGTAAACTGCTTCTGCTGACTCC	120	
Db	61	TCCAGCCCGGCTCCTCCTTGCTGTGACCTCCGAGTCCCTCAGTAAACTGCTTCTGCTGACTCC	120	
Qy	121	CATGTCTTTCACAGAGACTGAGCCAGTGCCTCAGAGGTTACCCCTTTGGCTTACACTGTTC	180	
Db	121	CATGTCTTTCACAGCAGACTGAGCCAGTGCCTCAGAGGTTACCCCTTTGGCTTACACTGTTC	180	
Qy	181	CTGCTGCTGCTGTGGACTTTAGCTTGGAGAAATGGAATCCACATGAGAGAGACCAG	240	
Db	181	CTGCTGCTGCTGTGGACTTTAGCTTGGAGAAATGGAATCCACATGAGAGAGACCAG	240	
Qy	241	GCACAGGACATTTCTGGGAGCAGTACCCCTTCTGCTGGAGGAGTGATGGCAGCAGCGGGA	300	
Db	241	GCACAGGACATTTCTGGGAGCAGTACCCCTTCTGCTGGAGGAGTGATGGCAGCAGCGGGA	300	
Qy	301	CAACTGGGACCCACTTGGCTCTCATCTCCTCTGGGGCAGCTTTCTGGACAGTCCGTC	360	
Db	301	CAACTGGGACCCACTTGGCTCTCATCTCCTCTGGGGCAGCTTTCTGGACAGTCCGTC	360	
Qy	361	CTCCTTGGGGCCCTGCACAGGCTCCTTGGAAACCCAGCTTCTCCACAGGCGAGGACCACA	420	
Db	361	CTCCTTGGGGCCCTGCACAGGCTCCTTGGAAACCCAGCTTCTCCACAGGCGAGGACCACA	420	
Qy	421	GCTCACAGGATGCCAAATGCCATCTTCCTGAGCTTTCACACCTGCTCCGAGGAAGGTG	480	

Db 421 GCTCACAGGATCCCAATGCAATCTTCTGAGCTTCCAAACACCTGCTCCGAGAAAGGTG 480

Qy 481 CGTTTCTGATGCTGTAGGAGGTCACACCTCTGCGTAGCG 522

Db 481 CGTTTCTGATGCTGTAGGAGGTCACACCTCTGCGTAGCG 522

RESULT 5

AR087571

LOCUS AR087571 1062 bp DNA PAT 07-SEP-2000

DEFINITION Sequence 18 from patent US 5989537.

ACCESSION AR087571

VERSION AR087571.1 GI:10014334

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1062)

AUTHORS Holly,R.D., Lok,S., Foster,D.C., Hagen,F.S., Kaushansky,K., Kuijper,J.L., Lofton-Day,C.E. and Oort,P.J.

TITLE Methods for stimulating granulocyte/macrophage lineage using thrombopoietin

JOURNAL Patent: US 5989537-A 18 23-NOV-1999;

FEATURES

source Location/Qualifiers

BASE COUNT 223 a 375 c 236 g 228 t

ORIGIN

Query Match 47.4%; Score 518.8; DB 6; Length 1062;

Best Local Similarity 99.6%; Pred. No. 5.8e-96;

Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGAGCTGACTGAATGCTCCTCGTGGTCATGCTTCTCCTAACTCAAGGCTAAGCGTG 50

Db 1 ATGGAGCTGACTGAATGCTCCTCGTGGTCATGCTTCTCCTAACTCAAGGCTAAGCGTG 50

Qy 61 TCCAGCCCGGCTCCTCCTGCTGTGACCTCCGAGTCCCTCAGTAACTGCTTCGTGACTCC 120

Db 61 TCCAGCCCGGCTCCTCCTGCTGTGACCTCCGAGTCCCTCAGTAACTGCTTCGTGACTCC 120

Qy 121 CATGCTCCTCACAGCAGACTGAGCCAGTCCGAGAGGTTTCAACCTTTGCTTACACCTGTC 180

Db 121 CATGCTCCTCACAGCAGACTGAGCCAGTCCGAGAGGTTTCAACCTTTGCTTACACCTGTC 180

Qy 181 CTGCTGCTGCTGTGAGCTTTAGCTTGGGAGATGGAAACCCAGATGGAGGAGACCAAG 240

Db 181 CTGCTGCTGCTGTGAGCTTTAGCTTGGGAGATGGAAACCCAGATGGAGGAGACCAAG 240

Qy 241 GCACAGGACATTTCTGGGAGCAGTGACCCCTCTGCTGGAGGAGTGATGGCAGCAGGGGA 300

Db 241 GCACAGGACATTTCTGGGAGCAGTGACCCCTCTGCTGGAGGAGTGATGGCAGCAGGGGA 300

Qy 301 CAACTGGGAGCCACTTGCCTCTCATCCCTCTGCGGAGCTTTCTGACAGGTCCTGTC 360

Db 301 CAACTGGGAGCCACTTGCCTCTCATCCCTCTGCGGAGCTTTCTGACAGGTCCTGTC 360

Qy 361 CTCTCTGGGGCCCTGAGAGCTTCTGCTGGAGATGGAAACCCAGATGGAGGAGACCAAG 420

Db 361 CTCTCTGGGGCCCTGAGAGCTTCTGCTGGAGATGGAAACCCAGATGGAGGAGACCAAG 420

Qy 421 GCACAGGACATTTCTGGGAGCAGTGACCCCTCTGCTGGAGGAGTGATGGCAGCAGGGGA 300

Db 421 GCACAGGACATTTCTGGGAGCAGTGACCCCTCTGCTGGAGGAGTGATGGCAGCAGGGGA 300

Qy 301 CAACTGGGAGCCACTTGCCTCTCATCCCTCTGCGGAGCTTTCTGAGCAGTCCGTC 360

Db 301 CAACTGGGAGCCACTTGCCTCTCATCCCTCTGCGGAGCTTTCTGAGCAGTCCGTC 360

Qy 361 CTCTCTGGGGCCCTGAGAGCTTCTGCTGGAGATGGAAACCCAGATGGAGGAGACCAAG 420

Db 361 CTCTCTGGGGCCCTGAGAGCTTCTGCTGGAGATGGAAACCCAGATGGAGGAGACCAAG 420

Qy 421 GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGAAAGGTG 480

Db 421 GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGAAAGGTG 480

Qy 481 CGTTTCTGATGCTTGTAGAGGTCACACCTCTGCGTAGCG 522

Db 481 CGTTTCTGATGCTTGTAGAGGTCACACCTCTGCGTAGCG 522

RESULT 6

I49760

LOCUS I49760 1062 bp DNA PAT 07-OCT-1997

DEFINITION Sequence 3 from patent US 5641655.

ACCESSION I49760

VERSION I49760.1 GI:2471980

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1062)

AUTHORS Foster,D.C., Helpel,M.D. and Holly,R.D.

TITLE Methods for producing thrombopoietin polypeptides using a mammalian tissue plasminogen activator secretory peptide

JOURNAL Patent: US 5641655-A 3 24-JUN-1997;

FEATURES

source Location/Qualifiers

BASE COUNT 223 a 375 c 236 g 228 t

ORIGIN

Query Match 47.4%; Score 518.8; DB 6; Length 1062;

Best Local Similarity 99.6%; Pred. No. 5.8e-96;

Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGAGCTGACTGAATGCTCCTCGTGGTCATGCTTCTCCTAACTCAAGGCTAAGCGTG 60

Db 1 ATGGAGCTGACTGAATGCTCCTCGTGGTCATGCTTCTCCTAACTCAAGGCTAAGCGTG 60

Qy 61 TCCAGCCCGGCTCCTCCTGCTGTGACCTCCGAGTCCCTCAGTAACTGCTTCGTGACTCC 120

Db 61 TCCAGCCCGGCTCCTCCTGCTGTGACCTCCGAGTCCCTCAGTAACTGCTTCGTGACTCC 120

Qy 121 CATGCTCCTCACAGCAGACTGAGCCAGTCCGAGAGGTTTCAACCTTTGCTTACACCTGTC 180

Db 121 CATGCTCCTCACAGCAGACTGAGCCAGTCCGAGAGGTTTCAACCTTTGCTTACACCTGTC 180

Qy 181 CTGCTGCTGCTGTGAGCTTTAGCTTGGGAGATGGAAACCCAGATGGAGGAGACCAAG 240

Db 181 CTGCTGCTGCTGTGAGCTTTAGCTTGGGAGATGGAAACCCAGATGGAGGAGACCAAG 240

Qy 241 GCACAGGACATTTCTGGGAGCAGTGACCCCTCTGCTGGAGGAGTGATGGCAGCAGGGGA 300

Db 241 GCACAGGACATTTCTGGGAGCAGTGACCCCTCTGCTGGAGGAGTGATGGCAGCAGGGGA 300

Qy 301 CAACTGGGAGCCACTTGCCTCTCATCCCTCTGCGGAGCTTTCTGACAGGTCCTGTC 360

Db 301 CAACTGGGAGCCACTTGCCTCTCATCCCTCTGCGGAGCTTTCTGACAGGTCCTGTC 360

Qy 361 CTCTCTGGGGCCCTGAGAGCTTCTGCTGGAGATGGAAACCCAGATGGAGGAGACCAAG 420

Db 361 CTCTCTGGGGCCCTGAGAGCTTCTGCTGGAGATGGAAACCCAGATGGAGGAGACCAAG 420

Qy 421 GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGAAAGGTG 480

Db 421 GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGAAAGGTG 480

Qy 481 CGTTTCTGATGCTTGTAGAGGTCACACCTCTGCGTAGCG 522

Db 481 CGTTTCTGATGCTTGTAGAGGTCACACCTCTGCGTAGCG 522

REFERENCE 1 (bases 1 to 1062)

LOCUS HUMTHROMB 1062 bp mRNA PRI 22-DEC-1994

DEFINITION Human thrombopoietin mRNA, complete cds.

ACCESSION L36052

VERSION L36052.1 GI:533216

KEYWORDS thrombopoietin.

SOURCE Homo sapiens cDNA to mRNA.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1062)

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Foster, D.C., Sprecher, C.A., Grant, F.J., Kramer, J.M., Kuijper, J.L., Holly, R.D., Whitmore, T.E., Heipel, M.D., Bell, L.A.N., Ching, A.F., McGrane, V., Hart, C., O'Hara, P.J., and Lok, S.

TITLE Human thrombopoietin: gene structure, cDNA sequence, expression, and chromosomal localization

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 91 (26), 13023-13027 (1994)

MEDLINE 95108091

FEATURES

source Location/Qualifiers

1..1062

/organism="Homo sapiens"

/db_xref="taxon:9606"

1..1062

/note="putative"

/codon_start=1

/product="thrombopoietin"

/protein_id="AAC37566.1"

/db_xref="GI:533217"

/translation="MELTELLLVMLLLTARLTLSSPAPACDLRLSKLLRDSHLVLRSLQCEPHPLPVLPAVDPSLGEWKTQMEETRAQDILGAPVTLLEGVMAARGQLGPTCLSLGQLSGQVRLGLGASLLGTOLPQPGRTTAHKDPNAIFLSQHLRLGKVRLMLVGGSTLCVRAPPTTAVPSRLVLTNLPNRTSGLLETFTASARTGSGLLKMQQGFRAKIPGLLNOTSRLDQIPGYLNRIHELLNGRLGFPSPSRRLTGLADISSGTSDTSLPNLQPGSPSPHTPTGQYTLFPLPPLTPVQVQLHLLPDPSPATPTP

TSPLNTSYTHSNLSQEG"

BASE COUNT 223 a 375 c 236 g 228 t

Query Match 47.4%; Score 518.8; DB 9; Length 1062;

Best Local Similarity 99.6%; Pred. No. 5.8e-96;

Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGCTGACTGAATTCCTCGTGCATGCTTCCTTAAGTGAAGCTAACGCTG 60

DB 1 ATGAGCTGACTGAATTCCTCGTGCATGCTTCCTTAAGTGAAGCTAACGCTG 60

QY 61 TCCAGCCGGCTCCTCGTGCATGCTTCCTTAAGTGAAGCTAACGCTG 120

DB 61 TCCAGCCGGCTCCTCGTGCATGCTTCCTTAAGTGAAGCTAACGCTG 120

QY 121 CATGCTCTTACAGCAGACTGAGCCAGTGCAGAGGTTTACCCCTTTCACCTGTC 180

DB 121 CATGCTCTTACAGCAGACTGAGCCAGTGCAGAGGTTTACCCCTTTCACCTGTC 180

QY 181 CTGCTGCTCTGCTGGAGAGTGAACCTTCTGCTGGAGGAGTGAAGGAGGAGG 240

DB 181 CTGCTGCTCTGCTGGAGAGTGAACCTTCTGCTGGAGGAGTGAAGGAGGAGG 240

QY 301 CAAGTGGAGCCACTTGCCTCTCATCCCTCTGCTGGAGGAGTGAAGGAGGAGG 360

DB 301 CAAGTGGAGCCACTTGCCTCTCATCCCTCTGCTGGAGGAGTGAAGGAGGAGG 360

QY 361 CTGCTGGGCGCTTGCAGAGCCCTTGGAAACCCAGCTTCTCCACAGGCGAGCACA 420

DB 361 CTGCTGGGCGCTTGCAGAGCCCTTGGAAACCCAGCTTCTCCACAGGCGAGCACA 420

QY 421 GCTCACAGGATCCCAATGCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAGG 480

DB 421 GCTCACAGGATCCCAATGCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAGG 480

QY 481 CGTTTCTGATGCTTGTAGAGGCTTCCACCTCTTGCCTGAGG 522

DB 481 CGTTTCTGATGCTTGTAGAGGCTTCCACCTCTTGCCTGAGG 522

RESULT 8

EI6669 1086 bp DNA PAT 28-JUL-1999

LOCUS

DEFINITION cDNA encoding thrombopoietin.

ACCESSION EI6669

VERSION E16669.1 GI:57111352

KEYWORDS JP 1998212243-A/3.

SOURCE Homo sapiens.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1086)

AUTHORS Kuroda, K.

TITLE COMPOSITION CONTAINING THROMBOPOIETIN PROTEIN AND USED FOR STIMULATING RECOVERY OF HEMOSTATIC THROMBUS-FORMING ABILITY OF

JOURNAL Patent: JP 1998212243-A 3 11-AUG-1998;

COMMENT KIRIN BREWERY CO LTD

OS Homo sapiens (human)

PN JP 1998212243-A/3

PD 11-AUG-1998

PF 31-JAN-1997 JP 1997019549

PI KURODA KENJI

PC A61K38/00,A01N1/02,C07K14/52//C12N15/09,C12P21/02.(C12P21/02,C12R1:19);

CC strandedness: Double;

CC topology: Linear;

CC hypothetical: No;

CC anti-sense: No;

Key Location/Qualifiers

FT source 1..1086

FT /organism="Homo sapiens"

FT /tissue_type="Liver"

FT 5'UTR 1..24

FT sig_peptide 25..87

FT mat_peptide 88..1086

FT /product="signal peptide of thrombopoietin"

FT CDS 25..1086

FT /product="thrombopoietin"

FEATURES

source Location/Qualifiers

1..1086

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 229 a 386 c 243 g 228 t

ORIGIN

Query Match 47.4%; Score 518.8; DB 6; Length 1086;

Best Local Similarity 99.8%; Pred. No. 5.8e-96;

Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGCTGACTGAATTCCTCGTGCATGCTTCCTTAAGTGAAGCTAACGCTG 60

DB 25 ATGAGCTGACTGAATTCCTCGTGCATGCTTCCTTAAGTGAAGCTAACGCTG 84

QY 61 TCCAGCCGGCTCCTCGTGCATGCTTCCTTAAGTGAAGCTAACGCTG 120

DB 85 TCCAGCCGGCTCCTCGTGCATGCTTCCTTAAGTGAAGCTAACGCTG 144

QY 121 CATGCTCTTACAGCAGACTGAGCCAGTGCAGAGGTTTACCCCTTTCACCTGTC 180

DB 145 CATGCTCTTACAGCAGACTGAGCCAGTGCAGAGGTTTACCCCTTTCACCTGTC 204

QY 181 CTGCTGCTCTGCTGGAGAGTGAACCTTCTGCTGGAGGAGTGAAGGAGGAGG 240

DB 205 CTGCTGCTCTGCTGGAGAGTGAACCTTCTGCTGGAGGAGTGAAGGAGGAGG 264

QY 241 GCACAGGACATTTGGGAGCAGTGACCTTCTCTGCTGGAGGAGTGAAGGAGGAGG 300

DB 265 GCACAGGACATTTGGGAGCAGTGACCTTCTCTGCTGGAGGAGTGAAGGAGGAGG 324

QY 301 CAAGTGGAGCCACTTGCCTCTCATCCCTCTGCTGGGCGAGCTTTCGACAGGTCCTC 360

DB 325 CAAGTGGGAGCCACTTGCCTCTCATCCCTCTGCTGGGCGAGCTTTCGACAGGTCCTC 384

QY 361 CTGCTTGGGCGCTTGCAGAGGCTTCTTGGAAACCCAGCTTCTTCCACAGGCGAGG 420

Dn	385	CTCCTTGGGCGCCCTCAGAGCCTCTCTTGGAACCCAGCTTCCTCCACAGGCAGGCCACA	444
Qy	421	GCTCACAAAGATCCCAATGCCACTTCTGTAGCTTCCAACACCTGCTCCGAGGAAAAGGTG	480
Dn	445	GCTCACAAAGATCCCAATGCCACTTCTGTAGCTTCCAACACCTGCTCCGAGGAAAAGGTG	504
Qy	481	CGTTTCTCATGCTGTGTAGGAGGTCCACCCTCTGCGTAGCGG	522
Dn	505	CGTTTCTCATGCTGTGTAGGAGGTCCACCCTCTGCGTAGCGG	546

RESULT 9

```

E12182
LOCUS      Human cDNA encoding thrombopoietin.          PAT      24-JUN-1998
DEFINITION
ACCESSION E12182
VERSION    1 GI:3251016
KEYWORDS   JP 1996277296-A/2.
SOURCE     Homo sapiens.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 1267)
AUTHORS    Miyazaki,H., Kato,H., Ogami,K., Iwamatsu,A., Akahori,H., Kuroki,R.,
            Shimizu,T. and Muto,T..
TITLE      PROTEIN HAVING TPO ACTIVITY
JOURNAL    Patent: JP 1996277296-A 2 22-OCT-1996;
            KIRIN BREWERY CO LTD
COMMENT    OS Homo sapiens (human)
            PN JP 1996277296-A/2
            PD 22-OCT-1996
            PF 14-FEB-1995 JP 1995161363
            PR 14-FEB-1994 JP 94P 35090, 25-MAR-1994 JP 94P 79842, PR
            01-JUN-1994 JP 94P 155126, 15-JUN-1994 JP 94P 167328, PR
            17-AUG-1994 JP 94P 227159, 17-DEC-1994 JP 94P 193169, PR
            01-NOV-1994 JP 94P 304167, 01-DEC-1994 JP 94P 298669, PR
            28-DEC-1994 JP 94P 341200
            PI MIYAZAKI HIROSHI, KATO HISASHI, OGAMI KINYA, IWAMATSU AKIHICO,
            PI AKAHORI HIRONORI, KUROKI RYOTA, SHIMIZU TOSHIYUKI, PI MUTO
            TAKANORI
            PC C07K14/52,C07H21/04,C12N15/09,C12P21/02//A61K38/00,(C12P21/02,
            PC C12R1:19),
            PC (C12P21/02,C12R1:91);
            CC strandedness: Double;
            CC topology: Linear;
            FH key Location/Qualifiers
            FT source 1..1267
            FT FT /organism='Homo sapiens'
            FT FT /tissue_type='liver'
            FT CDS 25..1086
            FT FT /product='thrombopoietin'
            FT sig_peptide 25..87.
            FT Location/Qualifiers
            FT 1..1267
            FT /organism='Homo sapiens'
            FT /db_xref='taxon:9606'
BASE COUNT 281 a 436 c 279 g 271 t
ORIGIN

```

Query Match 47.4%; Score 518.8; DB 6; Length 1267;

Best Local Similarity 99.6%; Pred. No. 5.6e-96;

Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	1	ATGGAGCTGACTGAATTGCTCTCGTGGTCATGCTTCTCTTAAGTCAAGGCTAACGGTTC	60
Dn	25	ATGGAGCTGACTGAATTGCTCTCGTGGTCATGCTTCTCTTAAGTCAAGGCTAACGGTTC	84
Qy	61	TTCAGCCCCGGCTCTCTCGTGGTGTGACCTCCGAGTCTCTAGTAAACTGCTTCGTGACTCC	120
Dn	85	TCAGCCCCGGCTCTCTCGTGGTGTGACCTCCGAGTCTCTAGTAAACTGCTTCGTGACTCC	144

Qy	121	CATGTCCTTCACAGCAGACTGAGCCAGTGGCCAGAGGTTCCACCTTTGCCCTACACCTGTC	180
Db	145	CATGTCCTTCACAGCAGACTGAGCCAGTGGCCAGAGGTTCCACCTTTGCCCTACACCTGTC	204
Qy	181	CTGCTGGCTGCTGCGACTTTAGCTTGGGAAGTGGAAACCCAGATGGAGGAGACCAAG	240
Db	205	CTGCTGGCTGCTGCGACTTTAGCTTGGGAAGTGGAAACCCAGATGGAGGAGACCAAG	264
Qy	241	GCACAGGACATTCTGGGAGCAGTACCCTTCTGCTGGAGGGAGTGATGGCAGCACGGGGA	300
Db	265	GCACAGGACATTCTGGGAGCAGTACCCTTCTGCTGGAGGGAGTGATGGCAGCACGGGGA	324
Qy	301	CAACTGGGACCCACTTGGCTCTCATCTCCTGGGGAGCTTTCTGGACAGGTCCGTCCTC	360
Db	325	CAACTGGGACCCACTTGGCTCTCATCTCCTGGGGAGCTTTCTGGACAGGTCCGTCCTC	384
Qy	361	CTCCTTGGGGCCCTGCAGAGGCTCTTGGAAACCAGCTTCTCCACAGGGAGGACCACA	420
Db	385	CTCCTTGGGGCCCTGCAGAGGCTCTTGGAAACCAGCTTCTCCACAGGGAGGACCACA	444
Qy	421	GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGTG	480
Db	445	GCTCACAGGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGTG	504
Qy	481	CGTTTCTGATGCTGTAGGAGGGTCCACCCTCTGCCGTACGG	522
Db	505	CGTTTCTGATGCTGTAGGAGGGTCCACCCTCTGCCGTACGG	546
RESULT 10			
E12214			
LOCUS	E12214	1267 bp	DNA
DEFINITION	Human cDNA encoding thrombopoietin.		
ACCESSION	E12214		
VERSION	E12214.1 GI:3251048		
KEYWORDS	JP 1996291196-A/1.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Miyazaki, H., Kato, H., Ogami, K., Iwamatsu, A., Akahori, H., Kuroki, R., Shimizu, T. and Muto, T.		
TITLE	PROTEIN HAVING TPO ACTIVITY		
JOURNAL	Patent: JP 1996291196-A 1 05-NOV-1996;		
COMMENT	KIRIN BREWERY CO LTD OS Homo sapiens (human) PN JP 1996291196-A/1 PD 05-NOV-1996 PF 14-FEB-1995 JP 1995355052 PR 14-FEB-1994 JP 94P 39090, 25-MAR-1994 JP 94P 79842, PR 01-JUN-1994 JP 94P 155126, 15-JUN-1994 JP 94P 167328, PR 17-AUG-1994 JP 94P 227159, 17-AUG-1994 JP 94P 193169, PR 01-NOV-1994 JP 94P 304167, 01-DEC-1994 JP 94P 298669, PR 28-DEC-1994 JP 94P 341200 PI MIYAZAKI HIROSHI, KATO HISASHI, OGAMI KINYA, IWAMATSU AKIHIKO PI AKAHORI HIROMORI, KUROKI RYOTA, SHIMIZU TOSHIYUKI, PI MUTO TAKANORI PC C07K14/53,C07K11/22,C12N15/09//A61K38/00,C12P21/08; CC strandedness: Double; CC topology: Linear; FH Key Location/Qualifiers FT source 1..1267 FT /organism='Homo sapiens' FT /tissue_type='liver' FT CDS 25..1086 FT /product='thrombopoietin' FT sig_peptide 25..87 FT mat_peptide 88..1083 FT /product='thrombopoietin'. FT Location/Qualifiers		
FEATURES			

source		1. .1267		/organism="Homo sapiens"		/db_xref="taxon:9606"		271 t	
BASE COUNT		281 a	436 c	279 g	271 t				
ORIGIN									
Query Match		47.4%;		Score 518.8;		DB 6;		Length 1267;	
Best Local Similarity		99.6%;		Pred. No. 5.6e-96;					
Matches 520;		Conservative 0;		Mismatches 2;		Indels 0;		Gaps 0;	
QY	1	ATGGAGCTGACTGAATTCCTCGTGGTCAATGCTTCTCCTAACTGAAGGCTAACGCTG	60						
Db	25	ATGGAGCTGACTGAATTCCTCGTGGTCAATGCTTCTCCTAACTGAAGGCTAACGCTG	84						
QY	61	TCCAGCCCGGCTCCTCGTGTGACCTCCGAGTCCAGTAACTGCTTCGTGACTCC	120						
Db	85	TCCAGCCCGGCTCCTCGTGTGACCTCCGAGTCCAGTAACTGCTTCGTGACTCC	144						
QY	121	CATGTCTTCACAGCAGACTGAGCCAGTGCAGAGGTTCCACCTTTGCTTACACCTGTC	180						
Db	145	CATGTCTTCACAGCAGACTGAGCCAGTGCAGAGGTTCCACCTTTGCTTACACCTGTC	204						
QY	181	CTGCTGCTGCTGGAGTCTTAGCTTGGGAGATGGAACCCAGATGGAGGAGACCAAG	240						
Db	205	CTGCTGCTGCTGGAGTCTTAGCTTGGGAGATGGAACCCAGATGGAGGAGACCAAG	264						
QY	241	GCACAGGACATTCCTGGAGCAGTACCCCTTCTGCTGGAGGAGTGGCAGCAGCGGGA	300						
Db	265	GCACAGGACATTCCTGGAGCAGTACCCCTTCTGCTGGAGGAGTGGCAGCAGCGGGA	324						
QY	301	CAACTGGGACCCACTTGCCTCTCATCCCTCTGCGGCAGCTTCTTGACAGGTCCTGTC	360						
Db	325	CAACTGGGACCCACTTGCCTCTCATCCCTCTGCGGCAGCTTCTTGACAGGTCCTGTC	384						
QY	361	CTCCTTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGCGAGACACA	420						
Db	385	CTCCTTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGCGAGACACA	444						
QY	421	GCTCACAAGATCCCAATGCCATCTTCCTGAGCTTCCAAACACCTGCTCCGAGGAAAGGTG	480						
Db	445	GCTCACAAGATCCCAATGCCATCTTCCTGAGCTTCCAAACACCTGCTCCGAGGAAAGGTG	504						
QY	481	CGTTTCTGTGCTGTAGGAGGTCACCCCTCTGCGTACGG	522						
Db	505	CGTTTCTGTGCTGTAGGAGGTCACCCCTCTGCGTACGG	546						
RESULT 11									
AR008878									
LOCUS		AR008878	1342 bp	DNA	PAT	04-DEC-1998			
DEFINITION		Sequence 1 from patent US 5756083.							
ACCESSION		AR008878							
VERSION		AR008878.1 GI:3967683							
KEYWORDS		Unknown.							
SOURCE		Unknown.							
ORGANISM		Unclassified.							
REFERENCE		1 (bases 1 to 1342)							
AUTHORS		Elliott, S.G.							
TITLE		Mpl ligand analogs							
JOURNAL		Patent: US 5756083-A 1 26-MAY-1998;							
FEATURES		Location/Qualifiers							
SOURCE		1. .1342							
BASE COUNT		299 a	454 c	293 g	296 t				
ORIGIN									
Query Match		47.4%;		Score 518.8;		DB 6;		Length 1342;	
Best Local Similarity		99.6%;		Pred. No. 5.5e-96;					
Matches 520;		Conservative 0;		Mismatches 2;		Indels 0;		Gaps 0;	
QY	1	ATGGAGCTGACTGAATTCCTCGTGGTCAATGCTTCTCCTAACTGAAGGCTAACGCTG	60						
Db	36	ATGGAGCTGACTGAATTCCTCGTGGTCAATGCTTCTCCTAACTGAAGGCTAACGCTG	95						
QY	61	TCCAGCCCGGCTCCTCGTGTGACCTCCGAGTCCAGTAACTGCTTCGTGACTCC	120						
Db	96	TCCAGCCCGGCTCCTCGTGTGACCTCCGAGTCCAGTAACTGCTTCGTGACTCC	155						
QY	121	CATGTCTTCACAGCAGACTGAGCCAGTGCAGAGGTTCCACCTTTGCTTACACCTGTC	180						
Db	156	CATGTCTTCACAGCAGACTGAGCCAGTGCAGAGGTTCCACCTTTGCTTACACCTGTC	215						
QY	181	CTGCTGCTGCTGGAGTCTTAGCTTGGGAGATGGAACCCAGATGGAGGAGACCAAG	240						
Db	216	CTGCTGCTGCTGGAGTCTTAGCTTGGGAGATGGAACCCAGATGGAGGAGACCAAG	275						
QY	241	GCACAGGACATTCCTGGAGCAGTACCCCTTCTGCTGGAGGAGTGGCAGCAGCGGGA	300						
Db	276	GCACAGGACATTCCTGGAGCAGTACCCCTTCTGCTGGAGGAGTGGCAGCAGCGGGA	335						
QY	301	CAACTGGGACCCACTTGCCTCTCATCCCTCTGCGGCAGCTTCTTGACAGGTCCTGTC	360						
Db	336	CAACTGGGACCCACTTGCCTCTCATCCCTCTGCGGCAGCTTCTTGACAGGTCCTGTC	395						
QY	361	CTCCTTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGCGAGACACA	420						
Db	396	CTCCTTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAGCTTCTCCACAGGCGAGACACA	455						
QY	421	GCTCACAAGATCCCAATGCCATCTTCCTGAGCTTCCAAACACCTGCTCCGAGGAAAGGTG	480						
Db	456	GCTCACAAGATCCCAATGCCATCTTCCTGAGCTTCCAAACACCTGCTCCGAGGAAAGGTG	515						
QY	481	CGTTTCTGTGCTGTAGGAGGTCACCCCTCTGCGTACGG	522						
Db	516	CGTTTCTGTGCTGTAGGAGGTCACCCCTCTGCGTACGG	557						
RESULT 12									
AR023468									
LOCUS		AR023468	1342 bp	DNA	PAT	05-DEC-1998			
DEFINITION		Sequence 24 from patent US 5795569.							
ACCESSION		AR023468							
VERSION		AR023468.1 GI:3976762							
KEYWORDS		Unknown.							
SOURCE		Unknown.							
ORGANISM		Unclassified.							
REFERENCE		1 (bases 1 to 1342)							
AUTHORS		Bartley, T.D., Bogenberger, J.M., Bosselman, R.A., Hunt, P., Kinstler, O.B. and Samal, B.B.							
TITLE		Mono-pegylated proteins that stimulate megakaryocyte growth and differentiation							
JOURNAL		Patent: US 5795569-A 24 18-AUG-1998;							
FEATURES		Location/Qualifiers							
SOURCE		1. .1342							
BASE COUNT		299 a	454 c	293 g	296 t				
ORIGIN									
Query Match		47.4%;		Score 518.8;		DB 6;		Length 1342;	
Best Local Similarity		99.6%;		Pred. No. 5.5e-96;					
Matches 520;		Conservative 0;		Mismatches 2;		Indels 0;		Gaps 0;	
QY	1	ATGGAGCTGACTGAATTCCTCGTGGTCAATGCTTCTCCTAACTGAAGGCTAACGCTG	60						
Db	36	ATGGAGCTGACTGAATTCCTCGTGGTCAATGCTTCTCCTAACTGAAGGCTAACGCTG	95						
QY	61	TCCAGCCCGGCTCCTCGTGTGACCTCCGAGTCCAGTAACTGCTTCGTGACTCC	120						
Db	96	TCCAGCCCGGCTCCTCGTGTGACCTCCGAGTCCAGTAACTGCTTCGTGACTCC	155						
QY	121	CATGTCTTCACAGCAGACTGAGCCAGTGCAGAGGTTCCACCTTTGCTTACACCTGTC	180						


```
Qy 481 CGTTTCCTGATGCTGTAGAGGGTCCACCTCTGCGTACGG 522
|||||
Db 516 CGTTTCCTGATGCTGTAGAGGGTCCACCTCTGCGTACGG 557
|||||

RESULT 15
D32047 1691 bp mRNA PRI 03-SEP-1997
DEFINITION Homo sapiens mRNA for thrombopoietin, complete cds.
ACCESSION D32047
VERSION D32047.1 GI:2351117
KEYWORDS thrombopoietin.
SOURCE Homo sapiens liver cdna to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Ogami, K.
Direct Submission
Submitted (08-JUL-1994) to the DDBJ/EMBL/GenBank databases. Kinya
Ogami, KIRIN Brewery Co., Ltd., Pharmaceutical Research Laboratory;
2-2 Soujamachi 1 chome, Maebashi, Gunma 371, Japan
(Tel:0272-54-8618, Fax:0272-52-2307)
2 (sites)
Kato, T., Ogami, K., Shimada, Y., Iwamatsu, A., Sohma, Y., Akahori, H.,
Horie, K., Kokubo, A., Kudo, Y., Maeda, E., Kobayashi, K., Ohashi, H.,
Ozawa, T., Inoue, H., Kawamura, K. and Miyazaki, H.
Purification and characterization of thrombopoietin
J. Biochem. 118 (1), 229-236 (1995)
96015174
Erratum: ([published erratum appears in J Biochem (Tokyo) 1996
Jan; 119(1):208])
FEATURES
source
1..1691
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="liver"
102..164
102..1163
/codon_start=1
/product="thrombopoietin"
/protein_id="BAA21930.1"
/db_xref="GI:2351118"
/translation="MELTELLVVMILLTARLTSSPPACDLRLVLSKLLRDSVHLH
SRLSQCEVHLEPTVLLPAVDFSLGKWKQMBETRAQDILGAVTLLEGVMAARGOL
GPTCLSLQLSGVRLGLGALQSLGTLQPPQGRTRAKDKNAIFLSQHLRLGRV
RFLMLVGGSTLTVRRAPPTAVPSRTSLVLTNLELNPRTSGLLFTNFTASARTGSGL
LKWQQGFRKIPGLNQTSLDQIFGLNRIHELNLNTRGLFPGPSRRTLGAPDISS
GTSDTGLPPLQPSPTHPPTGQYTLPLPPTLPTPVVQLHPLLPDPSAPTPTP
TSPLLNTSYTHSONLSQEG"
BASE COUNT 383 a 542 c 371 g 395 t
ORIGIN

Query Match 47.4%; Score 518.8; DB 9; Length 1691;
Best Local Similarity 99.6%; Pred. No. 5.1e-96;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGAGCTGACTGAATTGCTCCTCGTGGTCATGCTTCCTCAACTGCAAGGCTAACGCTG 60
|||||
Db 102 ATGGAGCTGACTGAATTGCTCCTCGTGGTCATGCTTCCTCAACTGCAAGGCTAACGCTG 161
|||||

Qy 61 TCCAGCCCGGCTCCTCGTGGTCATGCTTCGAGCTCCAGTAACTGCTTCGTCGACTCC 120
|||||
Db 162 TCCAGCCCGGCTCCTCGTGGTCATGCTTCGAGCTCCAGTAACTGCTTCGTCGACTCC 221
|||||

Qy 121 CATGTCCTTACAGCAGACTGAGCAGTGCAGGTTACACCTTTGCGCTACACCTGTC 180
|||||
Db 222 CATGTCCTTACAGCAGACTGAGCAGTGCAGGTTACACCTTTGCGCTACACCTGTC 281
|||||

Qy 181 CTGCTGCCTGCTGTGAGCTTTTAGCTTGGAGATGGAACCCAGATGGAGGAGACCAAG 240
|||||
```

Search completed: April 5, 2002, 10:29:40
Job time: 4773 sec

GenCore version 4.5
 Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2002, 10:33:11 ; Search time 188.75 Seconds
 (without alignments)
 4973.622 Million cell updates/sec

Title: US-09-680-514-8
 Perfect score: 1095
 Sequence: 1 ATGGAGCTGACTGAATTGCT.....TACGCCACCTTCCCGAGCCC 1095

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : N_Geneseq_1101.*
 1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.*
 2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.*
 3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.*
 4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.*
 5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.*
 6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.*
 7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.*
 8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.*
 9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.*
 10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.*
 11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.*
 12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.*
 13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.*
 14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.*
 15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.*
 16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.*
 17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.*
 18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.*
 19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.*
 20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.*
 21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.*
 22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1095	100.0	1095	17 AAT41788	Fusion peptide #3
2	1030.2	94.1	1083	17 AAT41787	Fusion peptide #2
3	987.4	90.2	1047	17 AAT41786	Fusion peptide #1
4	518.8	47.4	861	16 AAQ99552	Thrombopoietin cod
5	518.8	47.4	1062	16 AAT03941	Human thrombopoiet
6	518.8	47.4	1062	16 AAT04050	Sequence encoding
7	518.8	47.4	1062	17 AAT34852	Human thrombopoiet
8	518.8	47.4	1062	17 AAT37383	Human thrombopoiet
9	518.8	47.4	1062	17 AAT32591	Human thrombopoiet
10	518.8	47.4	1062	18 AAT85555	Human thrombopoiet
11	518.8	47.4	1062	19 AAV21696	Human thrombopoiet

12	518.8	47.4	1062	21 AAA51991	Human thrombopoiet
13	518.8	47.4	1086	17 AAT47958	Human thrombopoiet
14	518.8	47.4	1342	16 AAQ99704	Human MGFDF-1 and M
15	518.8	47.4	1342	17 AAT36657	Native human mpl 1
16	518.8	47.4	1342	17 AAT10025	Human MGFDF-1/2 cdn
17	518.8	47.4	1721	16 AAQ99554	Thrombopoietin cod
18	518.8	47.4	1721	21 AA240191	Human wild type th
19	518.8	47.4	1774	21 AAA47793	Human thrombopoiet
20	518.8	47.4	1795	16 AAQ94107	hML cDNA. Homo sa
21	518.8	47.4	1795	18 AAT64318	Human thrombopoiet
22	517.2	47.2	600	17 AAT36658	Truncated human mp
23	517.2	47.2	605	20 AAQ32813	Human truncated mp
24	517.2	47.2	1267	16 AAQ99553	Thrombopoietin cod
25	514	46.9	525	11 AAQ04482	Plasmid pASN6 enco
26	510.8	46.6	525	11 AAQ04484	Plasmid pASN145 en
27	506	46.2	525	11 AAQ04481	Plasmid pAS28 enco
28	495.8	45.3	1342	17 AAT33933	Human megakaryocyt
29	495.8	45.3	1342	19 AAV29068	Nucleotide sequenc
30	495.8	45.3	1342	20 AAQ32812	Human mpl ligand e
31	490.8	44.8	1062	19 AAV55365	Flt3L/IgG2b/G-CSF
32	490.4	44.8	1044	19 AAV55367	Flt3L/IgG2b/G-CSF
33	489.2	44.7	1002	19 AAV55401	Flt3L/GS/G-CSF c10
34	489.2	44.7	1002	19 AAV55403	Flt3L/GS/G-CSF c10
35	489.2	44.7	1003	19 AAV70366	Human flt3 ligand
36	489.2	44.7	1003	19 AAV70393	Nucleic acid seque
37	489.2	44.7	1003	19 AAV55110	G-CSF-Flt-3 ligand
38	489.2	44.7	1005	19 AAV70391	Nucleic acid seque
39	486.8	44.5	525	9 AAN80947	G-CSF gene isolate
40	486.8	44.5	1520	8 AAN70223	Plasmid pBRV2 inse
41	486.8	44.5	1520	8 AAN71320	Sequence encoding
42	486.8	44.5	1521	7 AAN60937	Plasmid pBRV2 inse
43	486.8	44.5	1525	9 AAN81478	Sequence encoding
44	486.8	44.5	1525	10 AAN91086	Plasmid pPL2 contg
45	486	44.4	921	16 AAQ97174	PMON13056 DNA enco

ALIGNMENTS

RESULT 1
 AAT41788
 ID AAT41788 standard; DNA; 1095 BP.
 XX
 AC AAT41788;
 XX
 DT 01-JUL-1997 (first entry)
 XX
 DE Fusion peptide #3 having G-CSF and TPO activity.
 XX
 KW Fusion protein; human granulocyte colony stimulating factor; hG-CSF;
 KW thrombopoietin; TPO; spacer peptide; blood platelet production;
 KW leukocyte production; anaemia; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT sig_peptide 1..63
 FT /*tag= a
 FT mat_peptide 64..1095
 FT /*tag= b
 XX
 PN WO9634016-A1.
 XX
 PD 31-OCT-1996.
 XX
 PF 26-APR-1996; 96WO-JP01157.
 XX
 PR 26-APR-1995; 95JP-0102625.
 XX
 PA (KYOW) KYOWA HAKKO KOGYO KK.
 XX
 PI Anazawa H, Konishi N, Shiotsu Y, Tamaoki T, Terasaki Y;
 PI Uchida K, Yamasaki M, Yamashita K, Yokoi H;

Qy	841	CAGGCCCTGGAGGGATCTCCCCCGAGTTGGGTGCCACCTTGGACACACTGCGAGCTGCAC	900
Db	829	cagccctggaaggatctccccccgagtTgggtcccaocttggacacactgcagctggac	888
Qy	901	GTGCCGACATTGCCACCCACCATCTCTGGCAGCAGATGGAGAACTGGGAATGGCCCTGCC	960
Db	889	gtgcgcgacttggccaccacatctggcagcagatggaagaactgggaatggccctgccc	948
Qy	961	CTGCAGCCCCACCCAGGGTGCACATGCCGGCCCTTCGCCTCTGTGCTTCAGAGGGCGGGCAGGA	1020
Db	949	ctgcagcccaccaccagagtgcataTgcgcggtcttcacgctctgtcttcacgcgcgcgacgaca	1008

Q1 1021 ggggtccragnctgctcccatctgcagagcttcctgaggtctcgaccgcgttctacgc 1068

Db 1069 cacttgccagccc 1083

|||||

RESULT 3

AAT41786

ID AAT41786 standard; DNA; 1047 BP.

XX AAT41786;

XX AC

DT 01-JUL-1997 (first entry)

XX

DE Fusion peptide #1 having G-CSF and TPO activity.

XX

KW Fusion protein; human granulocyte colony stimulating factor; hG-CSF;

KW thrombopoietin; TPO; spacer peptide; blood platelet production;

KW leukocyte production; anaemia; ds.

XX

OS Homo sapiens.

OS XX

Key Location/Qualifiers

1..63

FT sig_peptide /*tag= a

FT mat_peptide 64..1047

FT /*tag= b

XX

XX WO9634016-A1.

XX

XX 31-OCT-1996.

XX

XX 26-APR-1996; 96WO-JP01157.

XX

XX 26-APR-1995; 95JP-0102625.

XX

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX

XX Anazawa H, Konishi N, Shiotzu Y, Tamaaki T, Terasaki Y;

PI Uchida K, Yamasaki M, Yamashita K, Yokoi H;

XX

XX WPI; 1996-497573/49.

DR P-PSDB; AAW00377.

XX

XX Fusion peptide having G-CSF and with thrombopoietin activity -

PT optionally chemically modified with a poly(alkylene glycol, used for

PT treatment of anaemia

XX

XX Claim 3; Page 46-48; 71pp; Japanese.

XX

XX The sequences given in AAT41786-88 encode fusion proteins which are

CC composed of a peptide having human granulocyte colony stimulating

CC factor (hG-CSF) activity fused with a peptide having thrombopoietin

CC (TPO) activity, opt. via a spacer peptide. Peptides derived from

CC these by deletion, insertion or substitution of one or more amino

CC acid residues are included within the scope of the invention. The

CC fusion peptides stimulate blood platelet and leukocyte production

CC and are useful in the treatment of anaemia.

Db	913	ctgagccaccacccagggtgcacatgcgcgcttcgcctctgccttcacgcccggcgagga	972
Qy	1021	GGGTCCTAGTTCCTCCCATCTGCAGAGCTTCCTGGAGGTGTCGTACCGCGTTCTACGC	1080
Db	973	gggtcctagtgtgcctccatctgcagagcttcctgcgaggtgtcgtaccgcgttctacgc	1032
Qy	1081	CACCTGCCCCAGCCC	1095
Db	1033	cacctgcccagccc	1047
RESULT 4			
AAQ99552	AAQ99552 standard; cDNA to mRNA; 861 BP.		
XX	AC	AAQ99552;	
XX	DT	29-APR-1996 (first entry)	
XX	DE	Thrombopoietin coding sequence.	
XX	KW	Thrombopoietin; TPO; humoral factor; platelet; antibody; therapy; AIDS	
XX	KW	thrombocytopenia; hypoplastic anemia; thrombotic thrombocytopenia;	
XX	KW	disseminated intravascular coagulation syndrome; ds.	
XX	OS	Homo sapiens.	
XX	FH	Key	Location/Qualifiers
XX	FT	CDS	25..783
XX	FT		/*tag= a
XX	FT		/product= thrombopoietin
XX	FT	sig_peptide	25..87
XX	FT		/*tag= b
XX	FT	mat_peptide	88..783
XX	FT		/*tag= c
XX	XX	WO9521919-A2.	
XX	XX	17-AUG-1995.	
XX	XX	14-FEB-1995;	95WO-JP00208.
XX	XX	14-FEB-1994;	94JP-0039090.
XX	PR	14-MAR-1994;	94US-0212164.
XX	PR	25-MAR-1994;	94JP-0079842.
XX	PR	01-APR-1994;	94US-0221020.
XX	PR	01-JUN-1994;	94JP-0155126.
XX	PR	15-JUN-1994;	94JP-0167328.
XX	PR	20-JUL-1994;	94US-0278083.
XX	PR	17-AUG-1994;	94JP-0193169.
XX	PR	18-AUG-1994;	94JP-0193916.
XX	PR	11-OCT-1994;	94US-0320300.
XX	PR	01-NOV-1994;	94JP-0304167.
XX	PR	01-DEC-1994;	94JP-0298669.
XX	PR	22-DEC-1994;	94US-0361811.
XX	PR	28-DEC-1994;	94JP-0341200.
XX	PR	31-JAN-1995;	95US-0381478.
XX	PR	17-AUG-1994;	94JP-0227159.
XX	PA	(KIRI) KIRIN BREWERY KK.	
XX	PI	Akahori H, Iwamatsu A, Kato T, Kuroki R, Miyazaki H;	
XX	PI	Muto T, Ohgami K, Shimizu T;	
XX	XX	WPI; 1995-293120/38.	
XX	DR	P-PSDB; AAR81378.	
XX	XX	Thrombopoietin polypeptide which specifically stimulates or	
XX	PT	increases platelet production - useful in treatment of	
XX	PT	thrombocytopenia, also related DNA and vectors	
XX	PS	Claim 19; Page 337-338; 383pp; English.	

XX This sequence represents the coding sequence for a thrombopoietin (TPO)
CC protein. TPO is a humoral factor capable of promoting platelet
CC production. The encoded protein can be purified by using a gel affinity
CC column where Mpl has been coupled to a resin. AAQ99551 and AAQ99553
CC also encode TPO proteins. The DNA sequences are inserted into vectors
CC which are used to transform prokaryotic and eukaryotic host cells. Using
CC the proteins encoded by these sequences, and derivatives of them,
CC antibodies specifically immunoreactive with a TPO protein can be created.
CC The antibodies, DNA sequences and vectors are used to isolate the
CC protein sequences. The TPO proteins can then be used in the treatment
CC of platelet disorders. These include thrombocytopenia, hypoplastic
CC anaemia, AIDS, disseminated intravascular coagulation syndrome and
CC thrombotic thrombocytopenia.
XX
SQ Sequence 861 BP; 243 A; 257 C; 197 G; 164 T; 0 other;

Query Match 47.4%; Score 518.8; DB 16; Length 861;
Best Local Similarity 99.6%; Pred. No. 1.3e-112;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATGCTCTCGTGGTATGCTTCTCTCACTGCAAGGCTAACGCTG 60
DB 1 ATGGAGCTGACTGAATGCTCTCGTGGTATGCTTCTCTCACTGCAAGGCTAACGCTG 84
QY 61 TCCAGCCCGCTCTCTGCTGCTGAGCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
DB 85 TCCAGCCCGCTCTCTGCTGCTGAGCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 144
QY 121 CATGCTCTTCCAGCAGCAGTACGAGCTGCTGAGGCTTCCACCTTTGCTTACACCTGTC 180
DB 145 CATGCTCTTCCAGCAGCAGTACGAGCTGCTGAGGCTTCCACCTTTGCTTACACCTGTC 204
QY 181 CTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
DB 205 CTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 264
QY 241 GCACAGGACATCTGCGAGCAGTACCTTCTGCTGAGGAGTATGCGAGCAGCGGGA 300
DB 265 GCACAGGACATCTGCGAGCAGTACCTTCTGCTGAGGAGTATGCGAGCAGCGGGA 324
QY 301 CAACCTGGAGCCACTTGCCTCTCATCCCTCTGCTGAGGAGTATGCGAGCAGCGGTC 360
DB 325 CAACCTGGAGCCACTTGCCTCTCATCCCTCTGCTGAGGAGTATGCGAGCAGCGGTC 384
QY 361 CTCTTGGGCGCTGCGAGCCTCTTGGAGCCAGCTTCTTCCACAGGCGAGGACACA 420
DB 385 CTCTTGGGCGCTGCGAGCCTCTTGGAGCCAGCTTCTTCCACAGGCGAGGACACA 444
QY 421 GCTCACAAGGATCCCAATGCACTTCTGCTGAGCTTCCACACCTGCTCCGAGGAAGGTG 480
DB 445 GCTCACAAGGATCCCAATGCACTTCTGCTGAGCTTCCACACCTGCTCCGAGGAAGGTG 504
QY 481 CGTTTCTGATGTTGTAGGAGGTTCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 522
DB 505 CGTTTCTGATGTTGTAGGAGGTTCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 546

RESULT 5
AA03941
ID AA03941 standard; cDNA; 1062 BP.
XX
AC
XX
AC
AA03941;
DT 28-APR-1996 (first entry)

XX Human thrombopoietin coding sequence.
DE
XX
XX Thrombopoietin; erythropoiesis stimulator; treatment;
KW Thrombocytopenia; anaemia; ds.
XX
XX Homo sapiens.

XX Key Location/Qualifiers
FH 1..1062
FT /*tag= a
XX
XX W09521626-A1.
XX 17-AUG-1995.
XX
XX 09-FEB-1995; 95WO-US01829.
XX
XX 01-DEC-1994; 94US-0347748.
XX 14-FEB-1994; 94US-0196025.
XX 25-FEB-1994; 94US-0203197.
XX 21-MAR-1994; 94US-0215203.
XX 01-JUN-1994; 94US-0252491.
XX 09-AUG-1994; 94US-0288417.
XX 07-NOV-1994; 94US-0335566.
XX
XX (UNIW) UNIV WASHINGTON.
XX
XX Kaushansky K;
XX
XX WPI; 1995-292944/38.
XX P-PSDB; AAR82682.
XX
XX Stimulation of erythropoiesis using thrombopoietin and opt.
XX erythropoietin - for the treatment of thrombocytopenia and anaemia.
XX
XX Disclosure; Page 38-40; 66pp; English.
XX
XX This sequence encodes human thrombopoietin, which stimulates
XX erythropoiesis to produce an increase in proliferation or
XX differentiation of erythroid cells or to increase reticulocyte
XX counts at least 2-fold over baseline reticulocyte counts and,
XX optionally, platelet levels to at least 20000/cu mm. The protein
XX can be used in a composition, optionally with erythropoietin, for
XX use in the treatment of thrombocytopenia and anaemia, such as
XX that caused by destruction of haematopoietic cells in bone marrow.
XX In the treatment of cancer with chemotherapy and radiation, and in
XX pathological conditions such as myelodysplasia, AIDS, aplastic
XX anaemia, autoimmune disease or inflammatory disease.
XX
XX Sequence 1062 BP; 223 A; 375 C; 236 G; 228 T; 0 other;

Query Match 47.4%; Score 518.8; DB 16; Length 1062;
Best Local Similarity 99.6%; Pred. No. 1.4e-112;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATGCTCTCGTGGTATGCTTCTCTCACTGCAAGGCTAACGCTG 60
DB 1 ATGGAGCTGACTGAATGCTCTCGTGGTATGCTTCTCTCACTGCAAGGCTAACGCTG 60
QY 61 TCCAGCCCGCTCTCTGCTGCTGAGCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
DB 61 TCCAGCCCGCTCTCTGCTGCTGAGCTCCGAGTCTCAGTAACTGCTTCGTGACTCC 120
QY 121 CATGCTCTTCCAGCAGCAGTACGAGCTGCTGAGGCTTCCACCTTTGCTTACACCTGTC 180
DB 121 CATGCTCTTCCAGCAGCAGTACGAGCTGCTGAGGCTTCCACCTTTGCTTACACCTGTC 180
QY 181 CTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
DB 181 CTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 241 GCACAGGACATCTGCGAGCAGTACCTTCTGCTGAGGAGTATGCGAGCAGCGGGA 300
DB 241 GCACAGGACATCTGCGAGCAGTACCTTCTGCTGAGGAGTATGCGAGCAGCGGGA 300
QY 301 CAACCTGGAGCCACTTGCCTCTCATCCCTCTGCTGAGGAGTATGCGAGCAGCGGTC 360
DB 301 CAACCTGGAGCCACTTGCCTCTCATCCCTCTGCTGAGGAGTATGCGAGCAGCGGTC 360

CC protein that stimulates megakaryopoiesis and thrombocytopoiesis.
CC The cDNA can be incorporated into an expression vector and utilised
CC in the prodn. of low mol.wt. bioactive human thrombopoietin in
CC eukaryotic (partic. yeast or mammalian) host cells. The recombinant
CC thrombopoietin has an N-terminus at Ser-22 and a C-terminus between
CC Arg-185 and Asn-192 (esp. Arg-198 or Phe-207), and is useful for
CC increasing platelet prodn. in the treatment of thrombocytopenia or
CC for increasing proliferation of cells in bone marrow.
XX
SQ Sequence 1062 BP; 223 A; 375 C; 236 G; 228 T; 0 other;

Query Match 47.4%; Score 518.8; DB 17; Length 1062;
Best Local Similarity 99.6%; Pred. No. 1.4e-112;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGCTGACTGAATGCTCTCGTGGTCAATGCTTCTTCACTGAAGGCTTAACGCTG 60
|||||
DB 1 atgagctgactgaattgctctcggtgctatgcttctcttaactgaaggctaaacgtg 60
|||||
QY 61 TCAGAGCCGGCTCCTCTGCTTGTGACCTCCGAGTCTCAGTAACTGCTTCTGACTCC 120
|||||
DB 61 tcaagccgggctcctcgcttgagctccgagtcctcagtaaaactgcttgactcc 120
|||||
QY 121 CATGCTCTTCACAGCAGACTGAGCCAGTGCAGAGGTTTCCCTTTCCTACACCTGTC 180
|||||
DB 121 catgctcttcacagcagactgagccagtgccagaggttcacaccttggctacacctg 180
|||||
QY 181 CTGCTGCTCTGCTGACTTTAGCTTGGGAGATGGAACCCAGATGGAGGAGACCAAG 240
|||||
DB 181 ctgctgctgctgctgactttagcttggagaaatggaaacccagatggaggagaccaag 240
|||||
QY 241 GCACAGGACATTTCTGGAGCAGTGCCTTCTGCTGGAGGAGTGTGGCAGCAGCGGA 300
|||||
DB 241 gcacaggaacattctggagcagtgacccttctgctggagggagtgaaggagcagggga 300
|||||
QY 301 CAATGGGACCCACTTGTCTCTCATCCCTCTGGGGCAGCTTCTTGACAGAGTCCGTC 360
|||||
DB 301 caactgggacccacttgctctcatccctctgggagcttcttgagacaggtccgctc 360
|||||
QY 361 CTCTTGGGGCCCTGACAGCCCTCTTGGAAACCAGCTTCTTCCACAGGGCAGACCA 420
|||||
DB 361 ctctctggggccctgcagagcctccttggaaacccagcttctccacagggcaggaacca 420
|||||
QY 421 GCTCACAAAGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGGTG 480
|||||
DB 421 gctcaaaagatcccaatgccatcttctgagcttccaaacacctgctccgaggaaggtg 480
|||||
QY 481 CGTTTCTGATGCTTGTAGGAGGGTCCACCTCTGCGTACGG 522
|||||
DB 481 cgtttctgatgcttgtaggagggtccacctctcgctcagg 522
|||||

RESULT 8

AAAT37383
ID AAAT37383 standard; cDNA; 1062 BP.
AC
XC
AAAT37383;
XX
DT 25-NOV-1996 (first entry)
XX
XX Human thrombopoietin cDNA.
XX
KW Thrombopoietin; TPO; aplastic anaemia; congenital cytopaenia;
KW myelodisplastic syndrome; megakaryocytopoiesis; thrombocytopoiesis;
KW haematopoiesis; transgenic animal; ds.
XX
OS Homo sapiens.
XX
XX
XX Key Location/Qualifiers
FT sig_peptide 1..66
FT /*tag= a
FT mat_peptide 67..1059

FT /*tag= b

XX WO9620955-A1.
XX 11-JUL-1996.
XX 20-DEC-1995; 95WO-US16626.
XX 30-DEC-1994; 94US-0366859.
XX (ZYMO) ZYMOGENETICS INC.

XX Forstrom JW, Lofton-Day CE, Lok S;
XX WPI: 1996-333942/33.
XX P-PSDB; AAU03512.

XX New pure thrombopoietin free of low mol. wt. degradation prods. -
XX useful for treatment of aplastic anaemia, congenital cytopaenia, etc
XX Example 2; Page 63-65; 92pp; English.

XX A cDNA clone (AAT37383) codes for human thrombopoietin (TPO) (AAU03512).
CC It was isolated by PCR from human liver and kidney cDNA templates
CC using primers (see also AAT37385-88) derived from exon sequences of
CC a human genomic TPO clone (see also AAT37384) and from the conserved
CC 5' untranslated sequence of mouse TPO cDNA (AAT37382). The cDNA can
CC be used to produce recombinant mature TPO in transformed host cells.
CC Purified, homogeneous TPO is obtd. that is suitable for therapeutic
CC appln.

SQ Sequence 1062 BP; 223 A; 375 C; 236 G; 228 T; 0 other;

Query Match 47.4%; Score 518.8; DB 17; Length 1062;
Best Local Similarity 99.6%; Pred. No. 1.4e-112;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGCTGACTGAATGCTCTCGTGGTCAATGCTTCTTCACTGAAGGCTTAACGCTG 60
|||||
DB 1 atgagctgactgaattgctctcggtgctatgcttctcttaactgaaggctaaacgtg 60
|||||
QY 61 TCAGAGCCGGCTCCTCTGCTTGTGACCTCCGAGTCTCAGTAACTGCTTCTGACTCC 120
|||||
DB 61 tccagccgggctcctcgcttgtagctccgagtcctcagtaaaactgcttgactcc 120
|||||
QY 121 CATGCTCTTCACAGCAGACTGAGCCAGTGCAGAGGTTTCACTTTCCTACACCTGTC 180
|||||
DB 121 catgctcttcacagcagactgagccagtgccagaggttcacaccttggctacacctg 180
|||||
QY 181 CTGCTGCTCTGCTGGAGCTTTAGCTTGGGAGATGGAACCCAGATGGAGGAGACCAAG 240
|||||
DB 181 ctgctgctgctgctgagactttagcttggagaaatggaaacccagatggaggagaccaag 240
|||||
QY 241 GCACAGGACATTTCTGGAGCAGTGCCTTCTGCTGGAGGAGTGTGGCAGCAGCGGA 300
|||||
DB 241 gcacaggaacattctggagcagtgacccttctgctggagggagtgaaggagcagggga 300
|||||
QY 301 CAATGGGACCCACTTGTCTCTCATCCCTCTGGGGCAGCTTCTTGACAGAGTCCGTC 360
|||||
DB 301 caactgggacccacttgctctcatccctctgggagcttcttgagacaggtccgctc 360
|||||
QY 361 CTCTTGGGGCCCTGACAGCCCTCTTGGAAACCAGCTTCTTCCACAGGGCAGACCA 420
|||||
DB 361 ctctctggggccctgcagagcctccttggaaacccagcttctccacagggcaggaacca 420
|||||
QY 421 GCTCACAAAGATCCCAATGCCATCTTCTGAGCTTCCAAACACCTGCTCCGAGGAAAGGTG 480
|||||
DB 421 gctcaaaagatcccaatgccatcttctgagcttccaaacacctgctccgaggaaggtg 480
|||||
QY 481 CGTTTCTGATGCTTGTAGGAGGGTCCACCTCTGCGTACGG 522
|||||
DB 481 cgtttctgatgcttgtaggagggtccacctctcgctcagg 522
|||||

```

RESULT 9
AAT32591
ID AAT32591 standard; cDNA; 1062 BP.
XX
XX
AC AAT32591;
XX
XX 30-SEP-1996 (first entry)
DT
XX
DE Human thrombopoietin cDNA.
XX
XX Thrombopoietin; TPO; tissue plasminogen activator; cytokine;
KW protein secretion; signal peptide; thrombocytopenia; therapy; ds.
XX
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT sig_peptide 1..63
FT /*tag= a
FT mat_peptide 64..1059
FT /*tag= b
XX
XX WO9617067-A1.
PN
XX
XX 06-JUN-1996.
XX
XX 15-NOV-1995; 95WO-US14932.
XX
XX 30-NOV-1994; 94US-0347029.
XX
XX (ZYMO ) ZYMOGENETICS INC.
XX
XX Foster DC, Heipel MD, Holly RD;
XX
XX WPI; 1996-277784/28.
DR
DR P-PSDB; AAR99599.
XX
XX DNA construct encoding native mammalian tissue plasminogen activator
XX secretory peptide and thrombopoietin polypeptide - used to produce
XX thrombopoietin polypeptide(s) on a large scale and in cost effective
XX manner
XX
XX Disclosure; Page 41-43; 56pp; English.
XX
XX A cDNA clone (AAT32591) codes for human thrombopoietin (TPO) (AAR99599).
XX Novel DNA constructs code for a secretory peptide (AAR9600) fused to
XX the TPO polypeptide, with a cleavage site at the junction. Such
XX constructs can be expressed in transformed host cells, pref.
XX Pichia pastoris, Saccharomyces cerevisiae or rodent kidney cells,
XX for large-scale, cost-effective prodn. of human TPO, useful for the
XX treatment of thrombocytopenia.
XX
XX Sequence 1062 BP; 223 A; 375 C; 236 G; 228 T; 0 other;

Query Match 47.4%; Score 518.8; DB 17; Length 1062;
Best Local Similarity 99.6%; Pred. No. 1.4e-112; Indels 0; Gaps 0;
Matches 520; Conservative 0; Mismatches 2;

Qy 1 ATGGAGCTGACTGAATGTCTCTCGTGGTCATGCTTCTCTTAACAGGCTAACGCTG 60
Db 1 atggagctgactgaattgtctctcggtgcattgcttctcttaactgaaggtaacgctg 60
Qy 61 TCAGAGCCGCTCTCTGCTGTGTGACCTCGAGTCCAGTAACTGCTTGTGCTGACTCC 120
Db 61 tcagagccgctctctgtgtgacctcgagtcctcagtaactgcttgcgtgactcc 120
Qy 121 CATGTCTCTTCACAGCAGACTGAGCCAGTCCCGCAGAGGTTTCACCCCTTGCTACACTGTC 180
Db 121 catgtctcttcacagcagactgagccagtgccagaggttcaccccttgcctcacactgttc 180
Qy 181 CTGCTGCTGCTGTGGACTTTAGCTTTGGGAGAAATGGAAACCCAGATGGAGGAGACCAAG 240
Db 181 ctgctgctgctgtggaactttagctttgggagaaatggaaacccagatggaggagaccaag 240
Qy 241 GCACAGGACATTTCTGGAGCAGTACCCCTTCTGCTGAGGGAGTGTGGCAGCAGCGGGA 300
Db 241 gcacaggacattctggagcagtgacctctctgctgagggagtgatggcagcacgggga 300
Qy 301 CAACTGGGACCCACTTGCCTCTCATCCCTCTGGGCGAGCTTTCTGGACAGTCCGCTCTC 360
Db 301 caactgggacccacttgcctctcatccctctggggcagctttctggacaggtccgtctc 360
Qy 361 CTCTTTGGGGCCCTGCAGAGCCTCTCTTGGAACCCAGCTTCTCTCCACAGGCGAGACACA 420
Db 361 ctctttggggccctgcagagcctctcttggaaacccagcttctctccacagggcaggaacca 420
Qy 421 GCTCACAGGATCCCAATGCCATCTTCTTCCCTGAGCTTCCACACACTGCTCCGAGGAAAGTG 480
Db 421 gctcacaggatcccaatgcccatcttctctgagcttccacacctgctccgaggaaggtg 480
Qy 481 CGTTTCTCTGATGCTTGTAGGAGGTCACGCTCTGCTGCTACGG 522
Db 481 cgtttctctgatgcttgttagagggtccaccctctcgtcagg 522

RESULT 10
AAT85555
ID AAT85555 standard; cDNA; 1062 BP.
XX
XX
AC AAT85555;
XX
XX 05-NOV-1997 (first entry)
DT
XX
DE Human thrombopoietin cDNA.
XX
XX TPO; cancer; myeloid cell proliferation; expansion; bone marrow;
KW peripheral blood stem cell; chemotherapy; radiation therapy;
KW breast cancer; leukaemia; lymphoma; multiple myeloma; ds.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH CDS 1..1062
FT /*tag= a
FT mat_peptide 64..1059
FT /*tag= b
FT /*product= Thrombopoietin
XX
XX WO9640218-A1.
PN
XX
XX 19-DEC-1996.
PD
XX
XX 29-MAY-1996; 96WO-US07880.
PF
XX
XX 07-JUN-1995; 95US-0482212.
PR
XX
XX (ZYMO ) ZYMOGENETICS INC.
PA
XX
XX Fibbe WE, Grossmann A;
PI
XX
XX WPI; 1997-108626/10.
DR
XX
XX P-PSDB; AAW26841.
DR
XX
XX Increasing haematopoietic cells in patient - by admin. of stem cells
XX from donor previously treated with thrombopoietin, used in patients
XX who have received chemotherapy or radiation for cancer
XX
XX Disclosure; Page 17-19; 32pp; English.
XX
XX In a claimed method, thrombopoietin (TPO) (especially human TPO) is
XX administered to a donor to stimulate proliferation of myeloid cells.
XX Bone marrow or peripheral blood stem cells are collected from the
XX donor and administered to the recipient. The method is used for
XX preparing cells for (bone marrow) transplantation and for stimulating

```


PN W09628181-A1.
XX 19-SEP-1996.
XX 14-MAR-1996; 96WO-JP00635.
XX 15-MAR-1995; 95JP-0056248.
XX (KIRI) KIRIN BREWERY KK.
XX Otsuki N;
XX WPI; 1996-433541/43.
DR P-PSDB; AAW09314.
XX Stable thrombopoietin compositions contains TPO and an additive -
PT does not lose thrombopoietin activity on storage in a container
XX
XX Example 1; Page 26-28; 38pp; Japanese.
XX
XX Stable thrombopoietin (TPO) compsns. containing TPO and an additive,
CC pref. human serum albumin and/or gelatine, etc., are claimed. The
CC compsns. do not lose TPO activity on storage in a container as the
CC compsns. has reduced adsorption onto the container walls. The present
CC sequence encodes recombinant immature human TPO (1-332).
XX
SQ Sequence 1086 BP; 229 A; 386 C; 243 G; 228 T; 0 other;

Query Match 47.4%; Score 518.8; DB 17; Length 1086;
Best Local Similarity 99.6%; Pred. No. 1.4e-112;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGAGCTGACTGAATGCTCCCTCGTGTCTATCTCTTAAGTCAAGGCTTAACGCTG 60
DB 25 atggagctgactgaatgctccctcggtgtctatctcttaagtcgaaggttaacgctg 84
QY 61 TCCAGCCCGGCTCCTCGTGTGTGAGCTCCGAGTCTCAGTAACTGCTTGTGACTCC 120
DB 85 tccagcccggtcctcctggtgtgagctccgagtcctcagtaactgctgtgactcc 144
QY 121 CARGCTCTTACAGCAGCTGAGCCAGTGCCTGAGAGTTTACCCCTTTGCTTACACCTGTC 180
DB 145 catgtcttcacagcagactgagccagtgcagaggttcaccccttgcctacacctgtc 204
QY 181 CTGCTGCTGCTGGGAGCTTACTTGGGAGATGGAAACCCAGATGGAGGAGACCAAG 240
DB 205 ctgctgctgctgtggactttagcttggagatggaaacccagatggagagaccag 264
QY 241 GCACAGCACATTCTGGGAGCAGTGACCCCTTCTGCTGGAGGAGTGATGGCAGCACGGGGA 300
DB 265 gcacaggacattctgggagcagtgacccttctgctggagagtgatggcagcacgggga 324
QY 301 CACTGGGACCCACTTGCCTCTCATCCCTCCTGGGGCAGCTTCTTGACAGAGTCCGCTTC 360
DB 325 caactgggaccacttgcctctcatcctcctcctggtggcagcttcttgagacaggtccgtc 384
QY 361 CTCTTGGGGCCCTGACAGCCCTCTTGGAAACCCAGCTTCTCCACAGGCGACGACACA 420
DB 385 ctcttggggcccttgagagccctcttggaaacccagcttctccacagggcaggaccaca 444
QY 421 GCTCACAGGATCCCAATGCATCTTCTGCTGAGCTTCCACACCTGCTCCGAGGAAAGGTG 480
DB 445 gctcacagatcccaatgccattctctgagcttccacacactgctccgaggaagtg 504
QY 481 CGTTTCTGTATGCTTGTAGAGGGTCCACCCCTCTGCTGCTACGG 522
DB 505 cgtttctgtatgcttgttaggaggggtccacccctctgctgctcagg 546

RESULT 14

AAQ99704

ID AAQ99704 standard; cDNA; 1342 BP.

XX AAQ99704;
XX 01-FEB-1996 (first entry)
XX Human MGDF-1 and MGDF-2.
XX Mpl ligand; mammalian megakaryocyte growth promoting factor;
KW platelet producing factor; ss.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
CDS 36..1097
FT /*tag= a
XX
XX EP675201-A.
XX
XX 04-OCT-1995.
XX
XX 30-MAR-1995; 95EP-0104711.
XX
XX 30-NOV-1994; 94US-0347780.
XX 31-MAR-1994; 94US-0221768.
XX 31-MAY-1994; 94US-0252628.
XX 12-OCT-1994; 94US-0321488.
XX
XX (AMGE-) AMGEN INC.
XX
XX Bartley TD, Bogenberger JM, Bosselman RA, Hunt P;
PI Kinstler OB, Samal BB;
XX
XX WPI: 1995-338287/44.
XX P-PSDB; AAR80821, AAR80822, AAR80823, AAR80824, AAR80825, AAR80826.
XX
XX New human megakaryocyte growth and development factor - used to increase
PT the number of megakaryocytes or platelets in patients, e.g. for treating
PT thrombocytopenia
XX
XX Claim 17; Fig 11; 98pp; English.
XX
XX Mpl ligand is synonymous with MGDF. Mpl ligands are mammalian
CC megakaryocyte growth promoting and/or platelet producing factors.
CC MGDF-2 is a truncated portion of MGDF-1. It is hypothesized that
CC human MGDF is expressed in vivo as a substantially inactive or less
CC active precursor polypeptide that contains variable C-terminal AAs.
CC upon cleavage of the C-terminal AAs (as well as the signal peptide),
CC the processed form(s) of the molecule retain activity or become more
CC active. It is believed that MGDF-1 may require processing in order to
CC exhibit its activity. The fact that a truncated form MGDF-1 (i.e.
CC MGDF-2) is active supports this hypothesis. Various active molecules
CC that may result from truncations of the sequence set forth as
CC MGDF-1 (AAR80824) are given in AAR80822, AAR80823, & AAR80825.
CC is the signal peptide. The preferred truncation variants of MGDF-1
CC are any of those that have C-terminal truncations from AA 173-353
CC (along with cleavage of the signal peptide). The signal peptide
CC may have 23 AAs.
XX
XX Sequence 1342 BP; 299 A; 454 C; 293 G; 296 T; 0 other;

Query Match 47.4%; Score 518.8; DB 16; Length 1342;
Best Local Similarity 99.6%; Pred. No. 1.5e-112;
Matches 520; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGCTGACTGAATGCTCCCTCGTGTCTTCTCTTAAGTCAAGGCTTAACGCTG 60
DB 36 atggagctgactgaatgctccctcggtgtctctcttaagtcgaaggttaacgctg 95
QY 61 TCCAGCCCGGCTCCTCGTGTGTGAGCTCCGAGTCTCAGTAACTGCTTCTGACTCC 120
DB 96 tccagcccggtcctcctcggtgtgagctccgagtcctcagtaactgctgtgactcc 155

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2002, 09:47:02 ; Search time 1909.33 seconds
(without alignments)
6162.700 Million cell updates/sec

Title: US-09-680-514-8

Perfect score: 1095

Sequence: 1 ATGGAGCTGACTGAATGCT.....TAGGCACCTGCCAGGCC 1095

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_htc:*
10: gb_estl:*
11: gb_est2:*
12: gb_htc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	320.4	29.3	973	11	BI411128
2	312.8	28.6	598	11	BI411128
3	301.6	27.5	900	11	BI328788
4	300.6	27.5	507	10	BE485194
5	300.6	27.5	633	10	AI314551
6	299.6	27.4	609	10	AA538257
7	285.2	26.0	878	11	BF789308
8	227.8	20.8	464	10	BE754587
9	135.2	12.4	327	11	BF848766
10	135.2	12.3	449	13	AZ025993
11	106.2	9.7	419	13	AZ352219
12	86.4	7.9	708	13	AZ792988

13	77.8	7.1	752	13	AQ888740
14	65.6	6.0	589	10	BE377769
15	54.8	5.0	925	13	CNS0091P
16	52.4	4.8	976	11	BE823538
17	50.2	4.6	513	10	BE480590
18	49	4.5	925	13	CNS0091P
19	48.8	4.5	617	10	AI981598
20	44.6	4.1	702	11	BG433224
21	43.6	4.0	296	10	AI556398
22	43.2	3.9	461	10	BE598945
23	43.2	3.9	559	11	BE556176
24	42.8	3.9	427	11	BG704075
25	42.6	3.9	452	10	AA155632
26	42.6	3.9	899	13	CNS01VVT
27	42.4	3.9	1203	13	CNS015V4
28	42.2	3.9	598	10	AA115932
29	42	3.8	507	10	AI712717
30	42	3.8	553	10	AW253985
31	41.4	3.8	421	10	AA058743
32	41.4	3.8	478	11	BI345642
33	40.8	3.7	545	11	BF868325
34	40.6	3.7	447	11	BG604563
35	40.6	3.7	845	10	AL572931
36	40.6	3.7	932	13	CNS0072Q
37	40.6	3.7	1003	11	BI410408
38	40.6	3.7	1101	13	CNS05709
39	40.4	3.7	445	10	AW732737
40	40.4	3.7	458	10	AW273202
41	40.4	3.7	485	10	BE463718
42	40.4	3.7	490	10	AW410572
43	40.4	3.7	593	11	BG519711
44	40.2	3.7	864	13	CNS022LE
45	39.8	3.6	387	10	AW193322

ALIGNMENTS

RESULT. 1

BI411128	LOCUS	602962472F1	NCI_CGAP_Lu33	Mus musculus	cdna clone	IMAGE:5118185	5',
BI411128	DEFINITION	mRNA sequence.					
BI411128	ACCESSION	BI411128.1	GI:15172051				
BI411128	VERSION						
BI411128	KEYWORDS	EST.					
BI411128	SOURCE	house mouse.					
BI411128	ORGANISM	Mus musculus					
BI411128	REFERENCE	1	(bases 1 to 973)				
BI411128	AUTHORS	NIH-MGC	http://mgc.nci.nih.gov/				
BI411128	TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)					
BI411128	JOURNAL	Unpublished (1999)					
BI411128	COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabs-r@mail.nih.gov Tissue Procurement: Gilbert Smith, Ph.D. CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov Plate: LLAM11287 row: O column: 18 High quality sequence start: 32 High quality sequence stop: 912.					
BI411128	FEATURES	Location/Qualifiers					
BI411128	source	1..973					
BI411128	organism	"Mus musculus"					
BI411128	strain	"Czech II"					
BI411128	db_xref	"taxon:10090"					
BI411128	clone	"IMAGE:5118185"					

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 900)
REFERENCE NIH-MGC <http://mgc.ncl.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps@rmail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
plate: LLAM1338 row: g column: 19
High quality sequence stop: 808.

FEATURES
source
1..900
Location/Qualifiers
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5137578"
/clone_lib="NCI_CGAP_Li9"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: Sali; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.9 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library."
BASE COUNT 194 a 290 c 223 g 192 t 1 others
ORIGIN

Query Match 27.5%; Score 301.6; DB 11; Length 900;
Best Local Similarity 85.1%; Pred. No. 1.8e-57;
Matches 337; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 1 ATGAGCTGACTGAATGCTCTCGTGGTCAATGCTTCTCTAACTGCAAGGCTAACGCTG 60
|||||
DB 191 ATGAGCTGACTGATTGCTCTCGGCGGCATGCTTCTTGCAGTGGCAAGACTTA 250
QY 61 TCAGAGCCGCTCCTCTGCTGTGACCTCCGAGTCTCTAGTAACTGCTTCTGACTCC 120
|||||
DB 251 TCAGAGCCGCTGACTCTGCTGTGACCCAGACTCTCTAAATAAATGCTGCTGACTCC 310
QY 121 CATGCTCTTACAGACAGTACGACAGTGCAGAGGTTTACCTTTGCTTACACCTGTC 180
|||||
DB 311 CACCTCTTACAGACGACTGAGTGTGCTGACGCTGACCTTTGCTATCTCTGTT 370
QY 181 CTGCTGCTGCTGTGACTTTAGCTTGGGAGATGAAACCCAGATGGAGAGACCAAG 240
|||||
DB 371 CTGCTGCTGCTGTGACTTTAGCTTGGGAGATGAAACCCAGACGAGCAAG 430
QY 241 GCACAGGACATCTGGGACAGTACCTTCTGCTGGAGGAGTATGCGACGACGGGA 300
|||||
DB 431 GCACAGGACATCTAGGGGACAGTGTCTCTTCTACTGGAGGAGTATGCGACGACGAGGA 490
QY 301 CAACTGGGACCACTTGGCTCTCATCTCTCTGCGGAGCTTTCTGGACAGGTCCTGCTC 360
|||||
DB 491 CAGTTGGAACCTCTCTGCTCTCATCTCTCTGCGGAGCTTTCTGGCAGAGTTCTGCGCTC 550
QY 361 CTCCTTGGGGCCCTGCGAGACCTCTCTTGGAAACCCAG 396
|||||
DB 551 CTCTTGGGGCCCTGCGAGGCTCTCTCTAGGAACCCAG 586

RESULT 4
BE485194
LOCUS BE485194 507 bp mRNA EST 28-AUG-2000
DEFINITION 172130 BARC 580V Bos taurus cDNA 5', mRNA sequence.
ACCESSION BE485194
VERSION BE485194.1 GI:9604727
KEYWORDS EST.
SOURCE cow.

ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 507)
AUTHORS Sonstegard,T.S., Capuco,A.V., Van Tassell,C.P., Ashwell,M.S. and
Wells,K.D.
TITLE Mapping of Expressed Sequence Tags from a normalized bovine mammary
gland cDNA library
JOURNAL Unpublished (2000)
COMMENT Contact: Sonstegard TS
USDA, ARS, Beltsville Agricultural Research Center
Bldg. 200 Rm 2A, Beltsville, MD 20705, USA
Tel: 301 504 8416
Fax: 301 504 8414
Email: tads@psi.barc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.

PCR Primers
FORWARD: AGGAACACAGCTATGACCAT
BACKWARD: GTTTCAGCTCAGCAGC
Plate: 134 row: J column: 17
Seq primer: ATTAGGTGACACTATAG.

FEATURES
source
1..507
Location/Qualifiers
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="BARC 580V"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled mRNA isolated from mammary
tissues at eight physiological, developmental, and disease
states."
BASE COUNT 97 a 170 c 158 g 82 t
ORIGIN

Query Match 27.5%; Score 300.6; DB 10; Length 507;
Best Local Similarity 84.1%; Pred. No. 2.7e-57;
Matches 339; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 678 GTGTGTCACCTACAAGCTGTGCGACCCCGAGGAGCTGTGCTCGGACACTCTCTGGG 737
|||||
DB 104 GTGTGCGCGCCACAAGCTGTGCGACCCCGAGGAGCTGTGCTCGGACACTCTCTGGG 163
QY 738 CATCCCTCTGGCTCCCTGAGCAGCTGCCCCAGCCAGCCCTGCGAGCTGGCAGGCTGCTT 797
|||||
DB 164 CATCCCGCAGGCTCCCTAAGCAGCTGCTCCAGCCAGTCCCTGCGAGCTGACGCTGCCT 223
QY 798 GAGCCAACTCCATAGCGGCTTTTCTCTTACAGGGGCTCTTCCAGGCGCTCGAGGCGAT 857
|||||
DB 224 GAACCAACTACAGCGGGCTCTTCTCTTACAGGGGCTCTTCCAGGCGCTCGAGGCGAT 283
QY 858 CTCGCCCGAGTGGTGGTCCACCTTGGACACACTGAGCTGGAGCTGGCGAGCTTTGCCAC 917
DB 284 CTCGCCAGAGCTGGCGCCACCTTGGACACACTGAGCTGGAGCTGACCTGACTTTGCCAC 343
QY 918 CACCATCTGGCAGCAGATGGAAGACTGGGAATGGCCCTGCGCTGCGACCCACCCAGGG 977
DB 344 GAACATCTGGCTGAGATGGAGGACCTGGGGGGCGCCCCCGCTGTGTGACGCCACCCAGGG 403
QY 978 TGCCATGCGCGCTTCCGCTCTGCTTTCAGCGCGGCGAGGAGGCTCTAGTTGCCCTC 1037
DB 404 CGCATGCGCGACCTTCTTCTTACAGGGGCTCTTCCAGGCGCTTCCAGGAGGAGGCTCTGTTGCTTC 463
QY 1038 CCATCTGCGAGGCTTCTTGGAGGTGTGCTACCGGCTTCTTACGC 1080
DB 464 CCAGCTGCATCGTTCTTCTGAGCTGGCATACCGTGGCCTGCGC 506

RESULT 5

Db	358	CACCTCTTCACAGCCGACTGAGTCACTGTCCCGACGTCGACCCCTTGTCTATCCCTCTTT	417
QY	181	CTGCTGCTGCTGTGGACTTTAGCTTGGGAGATGGAAACCCAGATGGAGGACACCAAG	240
Db	418	CTGCTGCTGCTGTGGACTTTAGCTTGGGAGATGGAAACCCAGACGACAGCAAG	477
QY	241	GCACAGGACATTTCTGGGAGCAGTACCTTCTGCTGGAGGAGTGTATGGCAGCAGCGGA	300
Db	478	GCACAGGACATTTCTGGGAGCAGTACCTTCTGCTGGAGGAGTGTATGGCAGCAGCGGA	537
QY	301	CAACTGGGACCACTTGCCTCTCATCTCTCTGCTGGGAGCAGCTTCTGGACAGGTCGCTCTC	360
Db	538	CAGTTGGAACCCCTCTGCTCTCTCATCTCTCTGCTGGGAGCAGCTTCTGGGACAGTTCGCGCTC	597
QY	361	CTCTTGGGGCCCTGCAGAGCCTCTTGGAAACCCAG	396
Db	598	CTCTTGGGGCCCTGCAGGCGCTCTAGGAACCCAG	633

RESULT

6

AA538257

LOCUS

DEFINITION

vJ03807.r1 Barstead mouse pooled organs MPLRB4 Mus musculus cDNA
thrombopoietin mRNA, complete cds (MOUSE);, mRNA sequence.

ACCESSION

AA538257

VERSION

AA538257.1

KEYWORDS

GI:2284250

SOURCE

house mouse.

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 609)

AUTHORS

Marra,M., Hillier,L., Allien,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wyllie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.

TITLE

The WashU-HMI Mouse EST Project

JOURNAL

Unpublished (1996)

COMMENT

Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicinep
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:532844

FEATURES

Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 492.

source

Location/Qualifiers
1. .609
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:920628"
/clone_lib="Barstead mouse pooled organs MPLRB4"
/sex="mixed"
/tissue.type="pooled organs"
/dev_stage="7 day"
/lab_host="DH108"
/note="Organ: pooled; Vector: p7T3D-Pac (Pharmacia) with
a modified polylinker; Site1: EcoRI; Site2: NotI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer (5'
TGCTACGATCTGAAGTGGGAGCGCGCCCTTTTTTTTTTTTTTTTTTTT
3'); double-stranded cDNA was ligated to Eco RI adaptors
[GTTCGATTCGTTACC], digested with Not I and cloned into
the Not I and Eco RI sites of the modified p7T3 vector.
Library constructed by Bob Barstead."

BASE COUNT

126 a 204 c 157 g 122 t

ORIGIN

BE754587 464 bp mRNA EST 25-APR-2001
 2081371 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
 BE754587
 BE754587.1 GI:10168579
 EST.
 COW.
 Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 464)
 Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,
 Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett
 , G.L., Heaton, M.P., Iaegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G.,
 Perlman, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. and
 Keefe, J.W.
 Sequence evaluation of four pooled-tissue normalized bovine cDNA
 libraries and construction of a gene index for cattle
 Genome Res. 11 (4), 626-630 (2001)
 21180013
 Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called and alt-trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 18
 and -mtm4act 12 options.

library availability, please contact Pieter de Jong
(pieterdejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/cdb/bac_ends/mouse/bac_end_intro.html
Plate: 343 row: I column: 16
Seq primer: SP6
Class: BAC ends.

FEATURES

Location/Qualifiers
1..449
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-343I16"
/clone_lib="RPCI-23"
/sex="Female"
/lab_host="DH10B"

/note="Organ: Kidney/Brain; Vector: pBACE3.6; Site_1:
EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the pBACE3.6 vector at the
EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."
90 a 121 c 108 g 129 t 1 others

BASE COUNT

Query Match 12.3%; Score 135.2; DB 13; Length 449;
Best Local Similarity 84.4%; Pred. No. 2.7e-20;
Matches 152; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 228 GGAGGAGCAGGACAGACATCTGGAGCAGTGACCTTCTCTGGAGGAGTGAT 287
Db 126 GGAACAGAGCAGGACAGACATCTAGGGCAGTGCTTTTCTACTGGAGGAGTGAT 185

QY 288 GGCAGCAGGGGACACTGGACCCACTTGCCTCTCATCTCCTCTGGGCGAGCTTCTGG 347
Db 186 GGCAGCAGGAGCAGTGGAACCTTCTGCTCTATCTCTCTGGGAGGAGTCTGG 245

QY 348 ACAGGTGCGTCTCTCTGGGCGCTTGAGAGCCCTCTTGGAAACCCAGCTTCTCCACA 407
Db 246 GCAGGTTCGCTCTCTGGGCGCTTGAGAGCCCTCTAGGAACCCAGTAAAGTCCCA 305

RESULT 11

AZ352219/c
LOCUS 419 bp DNA GSS 29-SEP-2000
DEFINITION 1M0090P05R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0090P05 R, DNA sequence.

ACCESSION AZ352219
VERSION AZ352219.1 GI:10431456
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 419)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu

TITLE

JOURNAL

COMMENT

Insert Length: 10000 Std Error: 0.00
Plate: 0090 row: P column: 05
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 419.
Location/Qualifiers
1..419

FEATURES

Location/Qualifiers
1..419
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0090P05"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, P-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(<http://www.jax.org/resources/documents/dnares/>). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (gil4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid RL. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 100 a 98 c 124 g 97 t
ORIGIN

Query Match 9.7%; Score 106.2; DB 13; Length 419;
Best Local Similarity 79.2%; Pred. No. 8.1e-14;
Matches 126; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 360 CCTCTTGGGGCCCTCAGAGCCCTCTGGAACTTCTCTCCACAGGCGAGGCCAC 419
Db 292 CCTACTCACTGGCCTCAGGCCCATCTCTGCCCTCAGCTTCTCTACAGGGCAGGCCAC 233

QY 420 AGCTCAAGAGTCCCAATGCCATCTTCTGAGCTTCCACACCTCTCCGAGGAAGGT 479
Db 232 AGCTCAAGAGTCCCAATGCCATCTTCTGAGCTTCCACACCTCTCCGAGGAAGGT 173

QY 480 GCGTTTCTGTGCTGTAGGAGGGTCCACCTCTCGT 518
Db 172 GCGTTTCTGTGCTGTAGGAGGGTCCACCTCTCGT 134

RESULT 12

AZ792988
LOCUS 708 bp DNA GSS 16-FEB-2001
DEFINITION 2M0046105F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC2M0046105 F, DNA sequence.

ACCESSION AZ792988
VERSION AZ792988.1 GI:12937676
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 708)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

TITLE

